
WIDEOR (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Msrch_pp protein - protein database search, using Smith-Waterman algorithm
Date: Tue Mar 24 17:04:51 1998; MasPar time 3.63 Seconds
Similar output not generated. 358.098 Million cell updates/sec

Title: >US-08-236-918A-8
Description: (1-255) From 5674704.pep
Perfect Score: 1925
Sequence: 1 MGNSCYNIVATLLVLFNR.....QEDGCSGRPEEREGGCEL 255

Scoring table: PAM 150
Gap 11

Searched: 56404 seqs, 5096634 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:BACK1 2:51 3:52 4:53 5:54 6:55 7:56 8:57 9: PCT90
10: PCT91 11: PCT92 12: PCT93 13: PCT94 14: PCT95 15: PCT96

Statistics: Mean 30.414; Variance 146.494; scale 0.208

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Hit No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1925	100.0	255	15	PCT-US96-0	Sequence 8, Applicatio	4.99e-158
2	1925	100.0	255	7	US-08-236-	Sequence 8, Applicatio	4.99e-158
3	1116	58.0	256	7	US-08-236-	Sequence 6, Applicatio	2.44e-86
4	1116	58.0	256	15	PCT-US96-0	Sequence 2, Applicatio	2.44e-86
5	194	10.1	186	4	US-08-089-	Sequence 6, Applicatio	3.39e-07
6	183	9.5	326	5	US-08-292-	Sequence 4, Applicatio	2.46e-06
7	183	9.5	326	10	PCT-US91-0	Sequence 4, Applicatio	2.46e-06
8	174	9.0	325	10	PCT-US91-0	Sequence 2, Applicatio	1.23e-05
9	174	9.0	325	5	US-08-292-	Sequence 2, Applicatio	1.23e-05
10	153	8.5	355	5	US-08-292-	Sequence 6, Applicatio	8.63e-05
11	152	8.4	283	15	PCT-US96-1	Sequence 2, Applicatio	1.03e-04
12	140	7.3	206	5	US-08-097-	Sequence 7, Applicatio	4.69e-03
13	140	7.3	438	5	US-08-097-	Sequence 11, Applicati	4.69e-03
14	134	7.0	973	5	US-08-162-	Sequence 10, Applicati	1.30e-02
15	134	7.0	968	5	US-08-162-	Sequence 14, Applicati	1.30e-02
16	128	6.6	951	5	US-08-162-	Sequence 2, Applicatio	3.59e-02
17	127	6.6	970	14	PCT-US95-0	Sequence 11, Applicati	4.24e-02
18	123	6.4	461	1	5395760-2	Patent No. 5395760.	8.27e-02
19	123	6.4	461	7	US-08-385-	Sequence 2, Applicatio	8.27e-02
20	123	6.4	486	7	US-08-243-	Sequence 1, Applicatio	8.27e-02
21	123	6.4	518	7	US-08-385-	Sequence 4, Applicatio	8.27e-02
22	122	6.3	314	7	US-08-444-	Sequence 19, Applicati	9.76e-02

RESULT	1	ALIGNMENTS
ID	PCT-US96-03965-8	STANDARD: PRT: 255 AA.
AC	xxxxxx	
XX	01-JAN-1900	
DT	Sequence 8, Application PC/TUS9603965.	
XX		
CC	Sequence 8, Application PC/TUS9603965	
CC	GENERAL INFORMATION:	
CC	APPLICANT: Kwon, Byoung Se	
CC	APPLICANT: Kang, Chang-Yuill	
CC	TITLE OF INVENTION: Monoclonal antibody against human	
CC	NUMBER OF SEQUENCES: 10	
CC	CORRESPONDENCE ADDRESSES:	
CC	ADDRESS: Barnard, Brown & Michaels	
CC	STREET: 306 East State Street, Suite 220	
CC	CITY: Ithaca	
CC	STATE: NY	
CC	COUNTRY: USA	
CC	ZIP: 14850	
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE: Floppy disk	
CC	COMPUTER: IBM PC Compatible	
CC	OPERATING SYSTEM: PC-DOS/MS-DOS	
CC	SOFTWARE: Patentin Release #1.0, Version #1.30	
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER: PCT/US96/03965	
CC	FILING DATE:	
CC	CLASSIFICATION:	
CC	PRIOR APPLICATION DATA:	
CC	APPLICATION NUMBER: US 08/122,796	
CC	FILING DATE: 16-SEP-1993	
CC	PRIOR APPLICATION DATA:	
CC	APPLICATION NUMBER: US 08/012,269	
CC	FILING DATE: 01-FEB-1993	
CC	PRIOR APPLICATION DATA:	
CC	APPLICATION NUMBER: US 07/922,996	
CC	FILING DATE: 30-JUL-1992	
CC	PRIOR APPLICATION DATA:	
CC	APPLICATION NUMBER: US 07/267,577	
CC	FILING DATE: 07-NOV-1988	
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME: Michaels, Christopher A	

CC REGISTRATION NUMBER: 34,390
CC REFERENCE/DOCKET NUMBER: KMO5
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 607-273-1711
CC TELEFAX: 607-273-2609
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 255 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 255 AA: 27899 MW: 331875 CN;

Query Match 100.0%; Score 1925; DB 15; Length 255;
Best Local Similarity 100.0%; Pred. No. 4,99e-158;
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MGNSCYNIATLLVLFNFRTRSLDPCSNCPAGTFCDDNNRQICSPCPNSFSSAGGQR 60
OY 1 MGNSCYNIATLLVLFNFRTRSLDPCSNCPAGTFCDDNNRQICSPCPNSFSSAGGQR 60
QY 61 TCDICQCKGVRTRKRECSSTNAECDCTPGFHCGLGAGSMCEODCKQGOELTKKCKDC 120
QY 61 TCDICQCKGVRTRKRECSSTNAECDCTPGFHCGLGAGSMCEODCKQGOELTKKCKDC 120
Db 121 CGTFENDQKRGICRPWTNCSLDGKSVLVNGTKERDVYCGSPADLSPGASVTPPAPARE 180
OY 121 CGTFENDQKRGICRPWTNCSLDGKSVLVNGTKERDVYCGSPADLSPGASVTPPAPARE 180
Db 181 PGHSPOIISFFLALSTALFLFLFLLTFRFSVYKRGKKLLYIFKQPFMRPVQTQOEDG 240
OY 181 PGHSPOIISFFLALSTALFLFLFLLTFRFSVYKRGKKLLYIFKQPFMRPVQTQOEDG 240
Db 241 CSCRPEEEEGGCEL 255
OY 241 CSCRPEEEEGGCEL 255

RESULT 2
ID US-08-236-918A-8 STANDARD; PRT: 255 AA.
XX xxxxxx
DT 01-JAN-1900
DE Sequence 8, Application US/08236918A.
XX Sequence 8, Application US/08236918A
XX Patent No. 5674704
CC GENERAL INFORMATION:
CC APPLICANT: Alderson, Mark R.
CC APPLICANT: Goodwin, Raymond G.
CC APPLICANT: Smith, Craig A.
CC TITLE OF INVENTION: Cytokine Designated 4-1BB Ligand
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
CC STREET: 51 University Street
CC CITY: Seattle
CC STATE: Washington
CC COUNTRY: US
CC ZIP: 98101
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: Apple Power Macintosh
CC OPERATING SYSTEM: Apple 7.5.3
CC SOFTWARE: Microsoft Word, Version #6.0.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/236,918A
CC FILING DATE: 06-May-1994
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/060,843

CC FILING DATE: 07-May-1993
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Anderson, Kathryn A.
CC REGISTRATION NUMBER: 32,172
CC REFERENCE/DOCKET NUMBER: 2801-B
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (206) 587-0430
CC TELEFAX: (206) 233-0644
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 255 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 255 AA: 27899 MW: 331875 CN;

Query Match 100.0%; Score 1925; DB 7; Length 255;
Best Local Similarity 100.0%; Pred. No. 4,99e-158;
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MGNSCYNIATLLVLFNFRTRSLDPCSNCPAGTFCDDNNRQICSPCPNSFSSAGGQR 60
OY 1 MGNSCYNIATLLVLFNFRTRSLDPCSNCPAGTFCDDNNRQICSPCPNSFSSAGGQR 60
QY 61 TCDICQCKGVRTRKRECSSTNAECDCTPGFHCGLGAGSMCEODCKQGOELTKKCKDC 120
QY 61 TCDICQCKGVRTRKRECSSTNAECDCTPGFHCGLGAGSMCEODCKQGOELTKKCKDC 120
Db 121 CGTFENDQKRGICRPWTNCSLDGKSVLVNGTKERDVYCGSPADLSPGASVTPPAPARE 180
OY 121 CGTFENDQKRGICRPWTNCSLDGKSVLVNGTKERDVYCGSPADLSPGASVTPPAPARE 180
Db 181 PGHSPOIISFFLALSTALFLFLFLLTFRFSVYKRGKKLLYIFKQPFMRPVQTQOEDG 240
OY 181 PGHSPOIISFFLALSTALFLFLFLLTFRFSVYKRGKKLLYIFKQPFMRPVQTQOEDG 240
Db 241 CSCRPEEEEGGCEL 255
OY 241 CSCRPEEEEGGCEL 255

RESULT 3
ID US-08-236-918A-6 STANDARD; PRT: 256 AA.
XX xxxxxx
DT 01-JAN-1900
DE Sequence 6, Application US/08236918A.
XX Sequence 6, Application US/08236918A
XX Patent No. 5674704
CC GENERAL INFORMATION:
CC APPLICANT: Alderson, Mark R.
CC APPLICANT: Goodwin, Raymond G.
CC APPLICANT: Smith, Craig A.
CC TITLE OF INVENTION: Cytokine Designated 4-1BB Ligand
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
CC STREET: 51 University Street
CC CITY: Seattle
CC STATE: Washington
CC COUNTRY: US
CC ZIP: 98101
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: Apple Power Macintosh
CC OPERATING SYSTEM: Apple 7.5.3
CC SOFTWARE: Microsoft Word, Version #6.0.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/236,918A

```

FILING DATE: 06-May-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08,060,843
FILING DATE: 07-May-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Anderson, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2801-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 256 AA: 27598 MW; 347415 CN;

Query Match          58.0%; Score 1116; DB 7; Length 256;
Best Local Similarity 57.3%; Pred. No. 2,44e-86;
Matches 145; Conservative 49; Mismatches 55; Indels 4; Gaps 4;

Dd      1  MGNNCINNVYVLLVGCCEKAVQNQSCDNQPGTFCKRY-NPYCKACPSSTSSIGGP 59
Qy      1  MGNSCYNVAVLLVLYNLEFRTSRSLDDPCSNCPAGTFCDNNNQICSPCPNFSFSAAGOR 60
Dd      60  NCNICRVACAGYFRFKKFCSTHNADECIEGFHGLDPOCTCEKDCRPGDELTKQGC 119
Qy      61  TCDICRCKQGYFRFKRKECSTSNMECDCTPEFHCLGAGCSCEODCKQGLTKKGGKDC 120
Dd      120  SLGFENDQNGTGYCVRPWNCSLDGRSVLTKGTETKEDVYCGPVPVVSFSPSTTISVTPGCGP 179
Qy      121  CGFGEFNDQK-RGICRPWNCSLDKRSVLYNNTKRDVYCGPSPADLSPGASSVTPAPAR 179
Dd      180  G-GHSLQVLTFLALTS-ALLALILFTTLFSLVKWIRKRPPIHFKOPFKKTGGAQDEED 237
Qy      180  EPGHSPQIISFFLALTSALLFLFFLTLFRFSVYKRGKRKLITYFKOPFMRPVQTQDEED 239
Dd      238  ACSGRCPQEEEGG 250
Qy      240  GCSCRFPEEEEGG 252

RESULT      4
TIDY      PCT-US96-03965-2      STANDARD:      PRT:      256 AA.
XXXXXXXXXX
XX      01-JAN-1900
XX
DE      Sequence 2, Application PC/TUS9603965.
XX
XX      Sequence 2, Application PC/TUS9603965
XX
CC      GENERAL INFORMATION:
CC      APPLICANT: Kwon, Byoung Se
CC      APPLICANT: Kang, Chang-Yuill
CC      TITLE OF INVENTION: Monoclonal antibody against human
CC      TITLE OF INVENTION: receptor 4-1BB
CC      NUMBER OF SEQUENCES: 10
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Barnard, Brown & Michaels
CC      STREET: 306 East State Street, Suite 220
CC      CITY: Ithaca
CC      STATE: NY
CC      COUNTRY: USA
CC      ZIP: 14850
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC

```

```
CC      SOFTWARE: PatentIn Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: PCT/US96/03965
CC      FILING DATE:
CC      CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 08/122,796
CC      FILING DATE: 16-SEP-1993
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 08/012,269
CC      FILING DATE: 01-FEB-1993
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 07/922,996
CC      FILING DATE: 30-JUL-1992
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 07/267,577
CC      FILING DATE: 07-NOV-1988
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Michaels, Christopher A
CC      REGISTRATION NUMBER: 34,390
CC      REFERENCE/DOCKET NUMBER: KW05
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: 607-273-1711
CC      TELEFAX: 607-273-2609
CC      INFORMATION FOR SEQ. ID NO.: 2:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 256 amino acids
CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
SQ      SEQUENCE 256 AA; 27598 MW; 347415 CN;

Query Match          58.0%; Score 1116; DB 15; Length 256;
Best Local Similarity 57.3%; Pred. No. 2,44e-86;
Matches 145; Conservative 49; Mismatches 55; Indels 4; Gaps 4,

Db    1 MGNNCYNVYVLLVGCEKYGAQNQSCNDQPETFCKRT-NPYCKACPPSTSSIGOP 59
      ||||| :||| : :||| :||| :||| :||| :||| :||| :||| :|||
Qy    1 MGNSCNIVALLLVLFNERTRS:LDDPCSNCPACTFCDDNRNQICSPCPNSSFSAAGQR 60
      |::||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db    60 NCNLCRVACGFERRKRCSSTHNMEECSIEGPHLGPGCTRCEEDCRPGDELTKQGCKTC 119
      |::||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy    61 TCDICRCKGKGFERRKRKESSTSNAECCTPPEFHCLGAGCSNEODCKOGDELRKKGGKDC 120
      |::||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db    120 SLGTFDONGVGYCPWTWNCSLDRSVLKTGTTEKDVCYGPVVYSFSPTTISVTPEGGP 179
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy    121 CFGFFNQK-RGLGRPTWNCSDLKSRVLYVNKTRERDVYCGSPADLSFGASSVTPAPAR 179
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db    180 G-GHSLOYLTFLALTS-ALLALIIFTLLFSVLKWKIRKKEPHILFKOPFKRTTGAAQED 237
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy    180 EPGSPQLISFFALTSTALLFLFTLRPSVYKRGRRKLTYFKOPFMHPVQTGED 239
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db    238 ACSGRCPQEEEGG 250
Qy    240 GCSCRFPREEEGG 252

RESULT      5
ID   US-08-089-458B-6 STANDARD; PRT; 186 AA.
Ac     xxxxxx
Xx     01-JAN-1900
De     Sequence 6, Application US/08089458B.
Cc     Patent No. 5359039
Cc     GENERAL INFORMATION:
Cc     APPLICANT: Smith, Craig
Cc     APPLICANT: Goodwin, Raymond
Cc     TITLE OF INVENTION: Isolated Poxvirus A53R-Equivalent Tumor
Cc     TITLE OF INVENTION: Necrosis Factor Antagonists
```

```

CC NUMBER OF SEQUENCES: 7
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Patricia Anne Perkins, Immunex Corporation
CC STREET: 51 University Street
CC CITY: Seattle
CC STATE: Washington
CC COUNTRY: U.S.A.
CC ZIP: 98101
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: Apple Macintosh
CC OPERATING SYSTEM: Apple Operating System 7.1
CC SOFTWARE: Microsoft Word, Version #5.1a
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/089,458B
CC FILING DATE: 07/09/93
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Perkins, Patricia A.
CC REGISTRATION NUMBER: 34,693
CC REFERENCE/DOCKET NUMBER: 2608
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (206) 587-0430
CC TELEFAX: (206) 233-0644
CC TELEX: 756822
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 186 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE: 186 AA; 20372 MW; 170907 CN;

Query Match 10.1%; Score 194; DB 4; Length 186;
Best Local Similarity 31.3%; Pred. NO. 3.39e-07;
Matches 36; Conservative 24; Mismatches 42; Indels 13; Gaps 9;

Db 51 CNRCPGEPATIRCGSGSNTKRCRPPHTTYYTPNYSKCHQCKRCKPTGSDKYK-CTGT 109
Qy 28 CSNCPAGTF---CDNNNQICSPCPNPFSSAGG-ORTCDICRQC-KGVFRRKECSST 81
Db 110 ONSKSCJPGWFC-ATDSKTE-PCRDC-IPKRC-PCGYFGSIDELGNLICKS 159
Qy 82 SNAEDCTPGFHCLGAGSCWCEPDCKQGOELTKRGCKDC-CFGRFNQKRCICRP 135

RESULT 6
ID US-08-292-549-4 STANDARD; PRT; 326 AA.
AN xxxxxx
XX 01-JAN-1900
DE Sequence 4, Application US/08292549.
CC Sequence 4, Application US/08292549
CC Patent No. 5464938
CC GENERAL INFORMATION:
CC APPLICANT: Smith, Craig A.
CC APPLICANT: Goodwin, Raymond G.
CC TITLE OF INVENTION: Isolated Viral Protein TNF Antagonists
CC NUMBER OF SEQUENCES: 10
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunex Corporation
CC STREET: 51 University Street
CC CITY: Seattle
CC STATE: Washington
CC COUNTRY: USA
CC ZIP: 98101
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS

```

```
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/292,549
CC FILING DATE:
CC CLASSIFICATION: 530
CC PRIORITY INFORMATION DATA:
CC APPLICATION NUMBER: 07/963,330
CC FILING DATE: 10/19/92
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Perkins, Patricia A.
CC REGISTRATION NUMBER: 34,693
CC REFERENCE/DOCKET NUMBER: 2602-A
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (206) 587-0430
CC TELEFAX: (206) 233-0644
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 326 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 326 AA; 35208 MW; 586550 CN;
SQ
Query Match          9.5%; Score 183; DB 5; Length 326;
Best Local Similarity 31.5%; Pred. No.2,46e-06;
Matches      28; Conservative    20; Mismatches 32; Indels   9; Gaps
Db
40 CTCSPGVSAYSLRCGGPSTVCSPCKNETFTASTNHAPACVSCRGRCOTGLHSESOSCDKT 99
   |:|::||:| | :|::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Qy 28 CSNPATPF----CDNNRNQICSPCPNPSFSAGQG--RTDIDCR-QCKGVFRTRKESST 81
   ::::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 100 RDVYDCDSAGNYCLLKGQSGCRICAPKTK 128
   :|::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Qy 82 SNAECDCTPGFHCL-G-AGCSMKCEDCK 107
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
RESULT       7                STANDARD:           PRT;        326 AA.
ID            PCT-US91-02207-4
XX AC         xxxxxx
XX DT         01-JAN-1900
DE Sequence 4, Application PC/TUS9102207.
XX
CC Sequence 4, Application PC/TUS9102207
CC GENERAL INFORMATION:
CC APPLICANT: Smith, Craig A.
CC APPLICANT: Goodwin, Raymond G.
CC TITLE OF INVENTION: Isolated Viral Protein Cytokine Antagonists
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunex Corporation
CC STREET: 51 University Street
CC CITY: Seattle
CC STATE: Washington
CC COUNTRY: USA
CC ZIP: 98101
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.24
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US91/02207
CC FILING DATE: 19910329
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Wright, Christopher L.
CC REGISTRATION NUMBER: 31,680
CC REFERENCE/DOCKET NUMBER: 2602
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (206) 587-0430
```


[illegible]

```

CC TITLE OF INVENTION: Herpes Virus Entry Mediator
CC NUMBER OF SEQUENCES: 5
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Dressler, Goldsmith, Milnamow & Katz, Ltd.
CC STREET: 180 N. Stetson, Suite 4700
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: U.S.A.
CC ZIP: 60601
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/12374
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Northrup, Thomas E.
CC REGISTRATION NUMBER: 33,268
CC REFERENCE/DOCKET NUMBER: NOR3446P020PC
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (312) 616-5400
CC TELEFAX: (312) 616-5460
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 283 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 283 AA; 30519 MW; 429998 CN;
SQ
Query Match 8.4%; Score 162; DB 15; Length 283;
Best Local Similarity 30.7%; Pred. No. 1,03e-04;
Matches 23; Conservative 16; Mismatches 32; Indels 4; Gaps
Dd 67 CGELTGVCEPCPGPTIYIAHLNGISKLCQCOMCDPAGLPRATNCSRTENAVCGCSPGHF 126
Qy 37 CDNNRNQICSPCCPNSF-SSAGGGRCTDIDGRCR-GV-FRRKRCSSTSNACDCTGFGH 93
Dd 127 CIVQDGDHCA-ACRR 140
Qy 94 CLGAGCSMKCEDCKQ 108
RESULT 12
ID US-08-097-827-7 STANDARD; PRT; 206 AA.
XX AC xxxxxx
XX 01-JAN-1900
XX
Dd Sequence 7, Application US/08097827.
XX
CC Sequence 7, Application US/08097827
CC Patent No. 5457035
CC GENERAL INFORMATION:
CC APPLICANT: Baum, Peter
CC APPLICANT: Goodwin, Ray
CC APPLICANT: Fanslow, William
CC APPLICANT: Gayle, Richard
CC TITLE OF INVENTION: No. 5457035el cytokine which is a ligand for
CC TITLE OF INVENTION: OX40
CC NUMBER OF SEQUENCES: 13
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunex Corporation
CC STREET: 51 University Street
CC CITY: Seattle
CC STATE: WA
CC COUNTRY: USA
CC ZIP: 98101
CC

```


CC INFORMATION FOR SEQ ID NO: 10:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 973 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 973 AA; 107980 MW; 4944642 CN;

Query Match 7.0%; Score 134; DB 5; Length 973;
 Best Local Similarity 35.7%; Pred. No. 1.30e-02;
 Matches 25; Conservative 15; Mismatches 23; Indels 7; Gaps 5;

Db 250 AMKDTQCQACGPGTFRKSKQEGPCPPNSRTTAGATVC-ICR--SGFFRA--D-ADP 303
 QY 23 SLDDP-CSNCPACTPCDNNRNOICSPCPNFSAGQRTCDICRQCKGVFRTKCCSST 81

Db 304 ADSACTSVPS 313
 QY 82 SNAECDCTPG 91

IT 15
 ID US-08-162-809-14 STANDARD; PRT; 988 AA.

AC xxxxxx
 XX 01-JAN-1900

Sequence 14, Application US/08162809.

CC Sequence 14, Application US/08162809
 CC Patent No. 5457048

GENERAL INFORMATION:

CC APPLICANT: Pasquale, Elena B.
 CC APPLICANT: Sajjad, Fereydon G.
 CC TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
 CC TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
 CC NUMBER OF SEQUENCES: 26
 CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: CAMPBELL AND FLORES
 CC STREET: 4370 La Jolla Village Drive, Suite 700

CC CITY: San Diego
 CC STATE: California

CC COUNTRY: United States of America

CC ZIP: 92122

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/162,809

CC FILING DATE:

CC CLASSIFICATION: 514

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Campbell, Cathryn A.

CC REGISTRATION NUMBER: 31,815

CC REFERENCE/DOCKET NUMBER: P-LJ 9503

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (619) 535-9001

CC TELEFAX: (619) 535-8949

CC INFORMATION FOR SEQ ID NO: 14:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 988 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

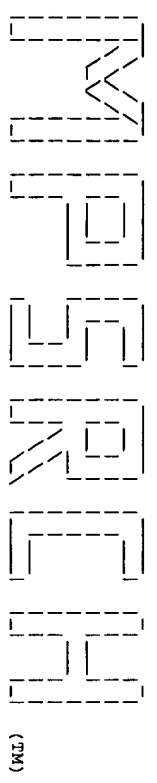
CC SEQUENCE 988 AA; 109578 MW; 5106875 CN;

CC Query Match 7.0%; Score 134; DB 5; Length 988;
 CC Best Local Similarity 35.7%; Pred. No. 1.30e-02;
 CC Matches 25; Conservative 15; Mismatches 23; Indels 7; Gaps 5;

Db 250 AMKDTQCQACGPGTFRKSKQEGPCPPNSRTTAGATVC-ICR--SGFFRA--D-ADP 303
 QY 23 SLDDP-CSNCPACTPCDNNRNOICSPCPNFSAGQRTCDICRQCKGVFRTKCCSST 81

Db 304 ADSACTSVPS 313
 QY 82 SNAECDCTPG 91

Search completed: Tue Mar 24 17:05:21 1998
 Job time : 30 secs.



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm
Date: Tue Mar 24 17:00:16 1998; Maspar time 7.76 Seconds
Database output not generated. 456,514 Million cell updates/sec

Title: >US-08-236-918A-8
Description: (1-255) from 5674704.pep
Perfect Score: 1925
Sequence: 1 MGNSCYNIVATLLVLNFER.....OEEDGSCRFPEEEGCGCEL 255

Scoring table: PAM 150
Gap 11

Searched: 111726 seqs, 13889129 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq30
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23

Statistics: Mean 32.592; Variance 147.726; scale 0.221
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1925	100.0	255	12	R64197	Human 4-1BB polypept1	7.29e-167
2	1925	100.0	255	13	R70977	H4-1BB receptor prote	7.29e-167
3	1897	98.5	255	18	W04174	Human receptor H4-1BB	3.21e-164
4	1871	97.2	255	13	R74087	Human receptor induce	9.13e-162
5	1116	58.0	255	19	W04173	Mouse receptor 4-1BB	8.27e-91
6	1116	58.0	255	12	R64199	Murine 4-1BB polypept	8.27e-91
7	1116	58.0	255	13	R70978	4-1BB receptor prote1	8.27e-91
8	199	10.3	260	4	R20814	T lymphocyte-specific	1.58e-07
9	199	10.3	260	18	R91441	Human CD27 antigen.	1.58e-07
10	194	10.1	186	11	R62655	Compox virus pst I/C1	4.09e-07
11	183	9.5	326	15	R85072	Myxoma virus T2 prote	3.26e-06
12	183	9.5	326	5	R27866	Myxoma virus T2 prote	3.26e-06
13	183	9.0	325	15	R85071	Shope fibroma virus T	1.75e-05
14	174	9.0	325	5	R27865	Rabbit fibroma virus	1.75e-05
15	163	8.5	355	15	R85073	Compox virus T2-equiv	1.34e-04
16	162	8.4	283	21	W05809	Human tumour necrosis	1.61e-04
17	162	8.4	283	20	W12659	Human herpes simplex	1.61e-04
18	158	8.2	197	23	W23220	Extracellular domain	3.35e-04
19	153	7.9	145	20	R39930	Osteoclastogenesis in	8.33e-04
20	147	7.6	417	5	R27736	Sequence transcribed	2.47e-03

21	143	7.4	401	20	R99931	Mutated OCIF, OCIF-C1	5.05e-03
22	141	7.3	154	20	R99929	Osteoclastogenesis in	7.22e-03
23	140	7.3	187	20	R99950	Mutated OCIF, OCIF-CB	8.63e-03
24	140	7.3	197	20	R99945	Mutated OCIF, OCIF-CD	8.63e-03
25	140	7.3	206	16	R18181	Mouse type-II membran	8.63e-03
26	140	7.3	272	20	R99944	Mutated OCIF, OCIF-CD	8.63e-03
27	141	7.3	277	13	R74737	ACT-4 cell surface re	7.22e-03
28	141	7.3	277	15	R79904	ACT-4-h1 receptor se	7.22e-03
29	140	7.3	321	20	R99949	Mutated OCIF, OCIF-CS	8.63e-03
30	140	7.3	326	20	R99940	Mutated OCIF, OCIF-DD	8.63e-03
31	140	7.3	327	20	R99941	Mutated OCIF, OCIF-DD	8.63e-03
32	140	7.3	351	20	R99943	Mutated OCIF, OCIF-CC	8.63e-03
33	140	7.3	360	20	R99938	Mutated OCIF, OCIF-DC	8.63e-03
34	140	7.3	380	20	R99924	Mature osteoclastogen	8.63e-03
35	140	7.3	390	20	R99357	Human tumour necrosis	8.63e-03
36	140	7.3	393	20	R99948	Mutated OCIF, OCIF-CB	8.63e-03
37	140	7.3	399	20	R99942	Mutated OCIF, OCIF-CL	8.63e-03
38	140	7.3	401	20	R99934	Mutated OCIF, OCIF-C2	8.63e-03
39	140	7.3	401	20	R99933	Mutated OCIF, OCIF-C2	8.63e-03
40	140	7.3	401	20	R99932	Mutated OCIF, OCIF-C2	8.63e-03
41	140	7.3	401	20	R99925	Full length osteoclas	8.63e-03
42	140	7.3	438	15	R81882	Plasmaid pDC406/OX40/F	8.63e-03
43	139	7.2	143	20	R99946	Mutated OCIF, OCIF-DC	1.03e-02
44	139	7.2	359	20	R99939	Mutated OCIF, OCIF-DC	1.03e-02
45	135	7.0	277	13	R76996	Deduced sequence enco	2.03e-02

ALIGNMENTS

RESULT 1
ID R64197 standard; Protein; 255 AA.
AC R64197;
DE 08-AUG-1995 (first entry)
DT Human 4-1BB polypeptide.
KW T-cell; lymphocyte; activation; tissue culture; clone; cell lines;
KW proliferation; stimulation; 4-1BB-L; ligand; receptor; cytokine.
OS Homo sapiens.
PN MO9426290-A.
PD 24-NOV-1994.
PE 06-MAY-1994; U05036.
PR 07-MAY-1993; US-060843.
PA (IMMUNEX) IMMUNEX CORP.
PI Alderson MR, Goodwin RG, Smith CA;
DR WPI; 95-022265/03.
DR N-PSDB; Q75424.
PT Cytokine, 4-1BB ligand (4-1BB-L) - binds to cell surface receptor
PT 4-1BB to transduce signal
PS Claim 39; Page 47-48; 65pp; English.
CC The 4-1BB polypeptide (receptor) and the 4-1BB-L (cytokine, see
CC Q75423) are useful in a pharmaceutical composition for stimulating
CC the immune system. The 4-1BB and 4-1BB-L polypeptides are also useful
CC for exploring mechanisms of T-cell activation, as they are expressed
CC on T lymphocytes. 4-1BB-L can be used as a tissue culture reagent for
CC in vitro cultivation of primary T-cells during the derivation of
CC clonal T-cell lines. It may also be used to stimulate proliferation
CC of activated T-cells, used in therapeutic procedures.
SQ Sequence 255 AA;

Query Match	100.0%;	Score 1925;	DB 12;	Length 255;
Best Local Similarity	100.0%;	Pred. No. 7.29e-167;		
Matches	255;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Db	1	mgnscynivatllvlnferisldqpsnppatfcdnmrngclspspnfsaagqr	60	
Oy	1	MGNSCYNIVATLLVLNFERISLDQPSNPPATFCDNMRNGCLSPSPNFSAGQR	60	
Db	61	tdclrcgkgyftrtrkecsstnaecdtpghclagagcmcegdckxgqeltkykdc	120	
Oy	61	TCDCRCCKGKGYFTRTRKECSSTNAECDTPGFHCLGAGACSMCEGDCXGQELTKGKDC	120	
Db	121	cfgtndgkrgricprwtcsldgksvlyngtkerdvvcgspadlspassvtpapare	180	
Oy	121	CFGTFNDGKRGRICPRWTCSLDGKSVLVNGTKERDVVCGSPADLSPASSVTPAPARE	180	

```

Db 181 pghspqisfflaltstallflfflltrfsvvkrgrklllyfkqpfmrpyqtgeedg 240
Qy 181 PGHSPQISFFLALTSTALLFLFFLLTRFSVVKRGKRKLLYFKQPFMRPYQTGEEDG 240
Db 241 cscrfeeeggccl 255
Qy 241 CSCRPEEEEGGCEL 255

RESULT 2
ID R70977 standard; Protein; 255 AA.
AC R70977;
DT 16-OCT-1995 (first entry)
DE H4-1BB receptor protein.
KW H4-1BB; receptor protein; immunosuppressive; autoimmune disease;
OS Homo sapiens.
PN W09507984-A.
PD 23-MAR-1995.
P 15-SEP-1994; U10457.
P 16-SEP-1993; US-122796.
P (INDV ) UNITV INDIANA FOUND.
PI Kwon BS;
DR WPI; 95-131352/17.
DR N-PSDB; 086126.
PT Novel cDNA encoding human receptor protein H4-1BB - useful to
PT produce the protein which is used to treat auto-immune disease
PT and facilitate organ transplantation
PS Claim 6; Fig.2; 36pp; English.
CC Human peripheral blood lymphocyte-derived cDNA was amplified by PCR
CC using probes based on the mouse receptor protein 4-1BB gene. The PCR
CC product was used to screen a cDNA library of activated human T-cells.
CC The isolated cDNA (086126), deposited as NRRL B21131, encoded the
CC human homolog, H4-1BB (R70977), of 4-1BB.
SQ Sequence 255 AA;

Query Match 100.0%; Score 1925; DB 13; Length 255;
Best Local Similarity 100.0%; Pred. No.7,29e-167;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mgnsynvatalllvlnfertrslqdpncpactfcdntrngicspcpnssfsaagqr 60
Qy 1 MGNSYNVATALLLVNFERTRSLQDPCSNCPAGFCDDNNRNGICSPCPNSSFSAAGQR 60
Db 61 tcdicrckgyvfrtrkccsstnaecdctpgfhclgagcmcegdckqgqeltkkqcdc 120
Qy 61 TCDICRCKGYVFRTRKCCSSTNAECDCTPGFHCLGAGCSMCEODCKQGOELTKKGCDC 120
Db 121 cffgtndkrgicrptwncslqgksvlyngtkerdvvgpspadlspgassvtpapare 180
Qy 121 CFFGTNDKRGICRPTWNCSLQGKSVLYNGTKERDVVCGPSPADLSPGASSVTPAPARE 180
Db 181 pghspqisfflaltstallflfflltrfsvvkrgrklllyfkqpfmrpyqtgeedg 240
Qy 181 PGHSPQISFFLALTSTALLFLFFLLTRFSVVKRGKRKLLYFKQPFMRPYQTGEEDG 240
Db 241 cscrfeeeggccl 255
Qy 241 CSCRPEEEEGGCEL 255

RESULT 3
ID W04174 standard; Protein; 255 AA.
AC W04174;
DT 12-DEC-1996 (first entry)
DE Human receptor H4-1BB.
KW Receptor 4-1BB; H4-1BB; monoclonal antibody; T-lymphocyte; T-cell;
KW B-lymphocyte; B-cell; immunostimulant; cancer; autoimmune disease;
KW graft rejection; therapy.
OS Homo sapiens.
PN W09629348-A1.
PD 26-SEP-1996.

```

```

PF 22-MAR-1996; U03965.
PR 23-MAR-1995; US-409851.
PA (INDV ) UNITV INDIANA FOUND.
PI Kang C; Kwon BS;
DR WPI; 96-443138/44.
DR N-PSDB; T39546.
PT Monoclonal antibody specific for human receptor protein 4-1BB - used
PT to enhance proliferation and activation of T-cells for treatment of
PT cancer and to inhibit specific ligand binding for treating
PT auto-immune diseases
PS Disclosure; Page 36-37; 48pp; English.
CC Novel human receptor protein H4-1BB (W04174) has the potential to
CC function as an accessory signaling molecule during T-cell activation
CC and proliferation. Its amino acid sequence was deduced from a cDNA
CC clone (T39546) isolated from activated human T-lymphocytes.
CC Recombinant H4-1BB can be produced in transfected host cells. It is
CC used to isolate receptor ligands, to stimulate B-cells expressing such
CC ligands and to block H4-1BB ligand binding. A fusion protein of
CC H4-1BB with human placental alkaline phosphatase can be used to modify
CC immune responses. A monoclonal antibody raised against an
CC immunopeptide (see also W04172) of H4-1BB is useful in the treatment
CC of cancer and autoimmune diseases.
SQ Sequence 255 AA;

Query Match 98.5%; Score 1897; DB 18; Length 255;
Best Local Similarity 99.6%; Pred. No.3,21e-164;
Matches 254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 mgnsynvatalllvlnfertrslqdpncpactfcdntrngicspcpnssfsaagqr 60
Qy 1 MGNSYNVATALLLVNFERTRSLQDPCSNCPAGFCDDNNRNGICSPCPNSSFSAAGQR 60
Db 61 tcdicrckgyvfrtrkccsstnaecdctpgfhclgagcmcegdckqgqeltkkqcdc 120
Qy 61 TCDICRCKGYVFRTRKCCSSTNAECDCTPGFHCLGAGCSMCEODCKQGOELTKKGCDC 120
Db 121 cffgtndkrgicrptwncslqgksvlyngtkerdvvgpspadlspgassvtpapare 180
Qy 121 CFFGTNDKRGICRPTWNCSLQGKSVLYNGTKERDVVCGPSPADLSPGASSVTPAPARE 180
Db 181 pghspqisfflaltstallflfflltrfsvvkrgrklllyfkqpfmrpyqtgeedg 240
Qy 181 PGHSPQISFFLALTSTALLFLFFLLTRFSVVKRGKRKLLYFKQPFMRPYQTGEEDG 240
Db 241 cscrfeeeggccl 255
Qy 241 CSCRPEEEEGGCEL 255

RESULT 4
ID R74087 standard; Protein; 255 AA.
AC R74087;
DT 21-JAN-1996 (first entry)
DE Human receptor induced by lymphocyte activation (ILA).
KW ILA; receptor inducible by lymphocyte activation; disease diagnosis;
KW antiinflammatory.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..17
FT /note- "signal peptide"
FT Domain 187..213
FT /note- "transmembrane domain"
FT MISC_difference 138
FT /note- "potential N-glycosylation site"
FT MISC_difference 149
FT /note- "potential N-glycosylation site"
FT MISC_difference 242
FT /note- "protein-kinase phosphorylation site"
FT MISC_difference 234
FT /note- "casein-kinase-II phosphorylation site"
FT MISC_difference 235
FT /note- "casein-kinase-II phosphorylation site"
FT Binding_site 241..244

```

FT /note="potential ligand binding site"
PN CA2108401-A.
PD 28-MAR-1995.
PF 14-OCT-1993; 108401.
PR 27-SEP-1993; US-127693.
PA (REGC) UNIV CALIFORNIA.
PI Lotz M, Schwarz H;
DR N-PSDB; R74087.
DR N-PSDB; R74087.
CC New receptor inducible by lymphocyte activation - used to develop
PT prods. for the diagnosis and treatment of inflammatory host defence
PT pathology.
PS Claim 55; Page 61; 91pp; English.
CC IIA may be used to identify a host defence inflammatory response in
CC body tissue. The IIA agents can be used to detect an IIA-mediated
CC pathology such as atherosclerosis, autoimmune disease (rheumatoid
CC arthritis), transplant rejection, pathogenic host defense responses
CC to microorganism and malignancy such as lung carcinoma.
SQ Sequence 255 AA;
Query Match 97.2%; Score 1871; DB 13; Length 255;
Best Local Similarity 98.4%; Pred. No. 9,13e-162;
Matches 251; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Db 1 mgncynvavtllvllvntfrrtsldgpcsnpcatfcdntrngdscppnfsaagqr 60
QY 1 MGNSCYNVAVTLLVLVNFERRSLDDPCSNCPAGTFCNNRNQICSPCPNFSFSSAGQR 60
Db 61 tcdlwcgkgyfrrfkrkccsstnsnecdtppgfwtlgaagcsmcedckogdeltrkckdc 120
QY 61 TCDICROCKGVRFRKRKCCSSTNSNECDTPGFHCLGAGCSMCEODCKOGDELTRKCKDC 120
Db 121 cftgfndqkrgicrptncslgkksvlnqtkerdvvcgspadlspassvtpapare 180
QY 121 CFTGFNDQKRGICRPTNCSLDGKSVLVNQTKEKRDVVCSPADLSPASSVTPAPARE 180
Db 181 pghspqiltsfllaltstallflflflltrfsvkrklllyfkqfmrpvqtgdeedg 240
QY 181 PGHSPQILTSFLLALTSTALLFLFLFLLTRFSVYKRRKLLYFKQFMRPVQTQEDD 240
Db 241 cscrfeeeggeeg 255
QY 241 CSCRFEEEGEGCEL 255
RESULT 5
ID W04173 standard; Protein; 256 AA.
AC W04173; 12-DEC-1996 (first entry)
DE Mouse receptor 4-1BB.
KW Receptor 4-1BB; H4-1BB; monoclonal antibody; T-lymphocyte; T-cell;
KW immunostimulant; cancer; autoimmune disease; graft rejection;
OS Mus sp.
PN W03623348-A1.
PD 26-SEP-1996.
PF 22-MAR-1996; U03965.
PR 23-MAR-1995; US-409851.
PA (INDV) UNIV INDIANA FOUND.
PI Kang C, Kwon BS;
DR WPI; 96-443138/44.
DR N-PSDB; T39541.
PT Monoclonal antibody specific for human receptor protein 4-1BB - used
PT to enhance proliferation and activation of T-cells for treatment of
PT cancer and to inhibit specific ligand binding for treating
PT auto-immune diseases
PS Disclosure; Page 32-34; 48pp; English.
CC Novel murine receptor protein 4-1BB (W04173) has the potential to
CC function as an accessory signaling molecule during T-cell activation
CC and proliferation. It may represent a cell surface molecule
CC involved in T-cell/APC interactions and may also act as a B-cell
CC costimulator. It is structurally related to members of the nerve
CC growth factor receptor superfamily. Its amino acid sequence was

CC deduced from an isolated cDNA clone (see also T39541). A human
CC homologue, H4-1BB (W04174), was identified and used to raise a
CC monoclonal antibody useful in cancer and autoimmune disease therapy.
SQ Sequence 256 AA;
Query Match 58.0%; Score 1116; DB 19; Length 256;
Best Local Similarity 57.3%; Pred. No. 8.27e-91;
Matches 145; Conservative 49; Mismatches 55; Indels 4; Gaps 4;
Db 1 mgncynvavtllvllvntfrrtsldgpcsnpcatfcdntrngdscppnfsaagqr 59
QY 1 MGNSCYNVAVTLLVLVNFERRSLDDPCSNCPAGTFCNNRNQICSPCPNFSFSSAGQR 60
Db 60 ncnicrvaagfrrfkfsssthaeccciegfhnlspqtrcckdcpqgdeltrkckdc 119
QY 61 TCDICROCKGVRFRKRKCCSSTNSNECDTPGFHCLGAGCSMCEODCKOGDELTRKCKDC 120
Db 120 slgftndqngtvcprptncslgkksvlnqtkerdvvcgspadlspassvtpapare 179
QY 121 CFTGFNDQKRGICRPTNCSLDGKSVLVNQTKEKRDVVCSPADLSPASSVTPAPARE 179
Db 180 g-gnslqvltrfllalts-allaiflltrfsvlkrkfkphfkqfkkttgaagdeed 237
QY 180 EPGHSPQILTSFLLALTSTALLFLFLFLLTRFSVYKRRKLLYFKQFMRPVQTQEDD 239
Db 238 aascrcpgeeg 250
QY 240 GSCRFEEEGEG 252
RESULT 6
ID R64199 standard; Protein; 256 AA.
AC R64199;
DE 08-AUG-1995 (first entry)
DE Murine 4-1BB polypeptide.
KW T-cell; lymphocyte; activation; tissue culture; clone; cell lines;
KW proliferation; stimulation; 4-1BB-L; ligand; receptor; cytokine.
OS Mus musculus.
PN W03426290-A.
PD 24-NOV-1994.
PF 06-MAY-1994; U05036.
PR 07-MAY-1993; US-060843.
PA (IMMUNEX) IMMUNEX CORP.
PI Alderson JM, Goodwin RG, Smith CA;
DR N-PSDB; Q75428.
DR N-PSDB; Q75428.
PT Cytokine, 4-1BB ligand (4-1BB-L) - binds to cell surface receptor
PT 4-1BB to transduce signal
PS Example 1; Page 44-45; 65pp; English.
CC The 4-1BB polypeptide (receptor) and the 4-1BB-L (cytokine, see
CC Q75422) are useful in a pharmaceutical composition for stimulating
CC the immune system. The 4-1BB and 4-1BB-L polypeptides are also useful
CC for exploring mechanisms of T-cell activation, as they are expressed
CC on T lymphocytes. 4-1BB-L can be used as a tissue culture reagent for
CC in vitro cultivation of primary T-cells during the derivation of
CC clonal T-cell lines. It may also be used to stimulate proliferation
CC of activated T-cells, used in therapeutic procedures.
SQ Sequence 256 AA;
Query Match 58.0%; Score 1116; DB 12; Length 256;
Best Local Similarity 57.3%; Pred. No. 8.27e-91;
Matches 145; Conservative 49; Mismatches 55; Indels 4; Gaps 4;
Db 1 mgncynvavtllvllvntfrrtsldgpcsnpcatfcdntrngdscppnfsaagqr 59
QY 1 MGNSCYNVAVTLLVLVNFERRSLDDPCSNCPAGTFCNNRNQICSPCPNFSFSSAGQR 60
Db 60 ncnicrvaagfrrfkfsssthaeccciegfhnlspqtrcckdcpqgdeltrkckdc 119
QY 61 TCDICROCKGVRFRKRKCCSSTNSNECDTPGFHCLGAGCSMCEODCKOGDELTRKCKDC 120
Db 120 slgftndqngtvcprptncslgkksvlnqtkerdvvcgspadlspassvtpapare 179
QY 121 CFTGFNDQKRGICRPTNCSLDGKSVLVNQTKEKRDVVCSPADLSPASSVTPAPARE 179

Qy	121	CGGTNDKQ-RCICRPWNCISDGKSVLVNGKEDVYCGGSPADLSGASSVPPAPAR	179
Db	180	9-gshslqvltlfialts-allaiaiffclisvlkwrkkfphkfgqfkttgaqeed	237
Qy	180	EPGHSPOIISFIALTSFALLFLPFLRLFSVYVARGRKLLYIRKQPFMNPVQTQED	239
Db	238	acsccrpggeeg 250	
Qy	240	GCSCRPFEEEGG 252	

RESULT 7			
ID	R70978	standard; Protein; 256 AA.	
AC	R70978;		
DT	16-OCT-1995	(first entry)	
DE	4-1BB receptor protein.		
KM	4-1BB; receptor protein; immunosuppressive; autoimmune disease;		
KW	organ transplantation; cell membrane ligand.		
OS	Mus sp.		
FT	Key	Location/Qualifiers	
PE	Peptide	1..23	
PI	/label- sig_peptide		
PD	MO9507984-A.		
PR	23-MAR-1995.		
PF	15-SEP-1994; U10457.		
PR	16-SEP-1993; U5122796.		
PA	(INDV) UNITV INDIANA FOUND.		
PI	Kwon BS;		
DR	WPI: 95-131352/17.		
DR	N-PSDB: Q86127.		
PT	Novel cDNA encoding human receptor protein H4-1BB - useful to		
PT	produce the protein which is used to treat auto-immune disease		
PS	and facilitate organ transplantation		
PT	Disclosure: Fig.1; 36pp; English.		
CC	cDNA encoding the human receptor protein H4-1BB (given in Q86126)		
CC	was isolated using PCR primers based on the homologous mouse		
CC	4-1BB gene (Q86127) encoding mouse receptor protein (R70978).		
CC	Sequence 256 AA;		

Query Match	58.0%;	Score 1116;	DB 13;	Length 256;
Best Local Similarity	57.3%;	Pred. No. 8, 27e-91;		
Matches 145;	Conservative 49;	Mismatches 55;	Indels 4;	Gaps 4

Db	1	mgancyvvvvi11lygcekyavagvnsdcngcpgtfcfky-npyvkscpstsfsgyqp	59
Qy	1	MNSCNVIAVATLLVLNFERSTRLODPCSNCPAGTCDNNRNQICSPCPMNSFSAGOR	60
Db	60	nonictvcegyfirfkfststmaacecelegfhcigpqtctcekdcrpgeltqgctk	119
Qy	61	TCDDICRQCGVFRTRKESSTSNACDCTPGFHCIGAGCSMCEDCKGQGLTRKGCDC	120
Db	120	slgtfdngqgtvccpwtnccsidgrsvlktgttekdvcgppvysfssstislvtpegp	179
Qy	121	CGGTNDKQ-RCICRPWNCISDGKSVLVNGKEDVYCGGSPADLSGASSVPPAPAR	179
Db	180	9-gshslqvltlfialts-allaiaiffclisvlkwrkkfphkfgqfkttgaqeed	237
Qy	180	EPGHSPOIISFIALTSFALLFLPFLRLFSVYVARGRKLLYIRKQPFMNPVQTQED	239
Db	238	acsccrpggeeg 250	
Qy	240	GCSCRPFEEEGG 252	

RESULT 8			
ID	R20814	standard; Protein; 260 AA.	
AC	R20814;		
DT	21-MAY-1992	(first entry)	
DE	T lymphocyte-specific CD27 Antigen.		
KM	Rapid immunoselection cloning technique; cell surface antigen;		
KW	carcinoma antigen CD40.		
OS	Homo sapiens.		
FT	Key	Location/Qualifiers	

FT	Peptide	1..20
FT	/label= signal	
FT	Protein	21...260
FT	/label= CD27_antigen	
FT	Domain	21..191
FT	/label= extracellular	
FT	Region	192..211
FT	/label= transmembrane	
FT	Domain	212..260
FT	/label= cytoplasmic	
PN	MO9201049-A.	
PD	23-JAN-1992.	
PF	15-JUL-1990.	U04986.
PR	13-JUL-1990.	US-553759.
PA	(GENO-) GEN HOSPITAL CORP.	
PI	Seed B, Aruffo A, Amiot M;	
DR	WPI: 92-056864/07.	
DR	N-PSDB: Q21183.	
PT	Immune CD33 cell surface antigen and DNA encoding it - for	
PT	immuno-therapy and diagnosis of haematopoietic neoplasms, etc.	
PS	Example 13; Page 103; 160pp; English.	
CC	The amino acid sequence of CD27 was deduced from a cDNA clone	
CC	obtained from human T lymphocyte cDNA transferred into COS cells	
CC	and immunoselected using the Mabs OKT18a and CLB-9F4 (see e.g.	
CC	Q21164 for description of the rapid immunoselection cloning method).	
CC	The CD27 polypeptide demonstrates the typical features of a type I	
CC	integral membrane protein. The hydrophobic membrane spanning region	
CC	is followed by a cytoplasmic domain beginning with a positively	
CC	charged stop transfer sequence. The deduced CD27 sequence is highly	
CC	homologous to the B lymphocyte and carcinoma antigen CD40 over its	
CC	entire length. CD27 is also homologous to the NGF receptor over the	
CC	extracellular and transmembrane domains. There is a cysteine- and	
CC	histidine-rich region in the extracellular domain which could	
CC	potentially form "zinc fingers", followed by a serine, threonine	
CC	and proline rich membrane proximal domain.	
SQ	Sequence	260 AA;

Query Match 10.3%; Score 199; DB 4; Length 260;
 Best Local Similarity 38.6%; Pred. No. 1.58e-07;
 Matches 27; Conservative 14; Mismatches 25; Indels 4; Gaps 4;

Db	53	cdqhnrkaqcddcpibpyvgsfspdhttrphcscrhnsqllvrn-clttanaecacrnwg	111
QY	37	CDNNNNQI-CSPCPPN-SFSSAGGQRT-CDICRCKGCVFRTKRCSSYNAECDCTPGFH	93

Db	112	crdckctcd 121
QY	94	CLGAGCSMCE 103

RESULT 9

ID	R91441	standard; Protein; 260 AA.
AC	R91441:	
DT	31-OCT-1996	(first entry)
DE	Human CD27 antigen.	
KW	Cell surface antigen; cloning; immunoselection; immunotherapy;	
KW	therapy; diagnosis; vector; CD27 antigen; T-lymphocyte;	
KW	autoimmune disease.	
OS	Homo sapiens.	
FT	Key	Location/Qualifiers
FT	Peptide	1..20
FT	/label= Sig_peptide	
FT	Domain	21..91
FT	/label= Extracellular_domain	
FT	Domain	92..212
FT	/label= Transmembrane_domain	
FT	Domain	213..260
FT	/label= Cytoplasmic_domain	
PN	US506126-A.	
PD	09-APR-1986.	
PF	25-FEB-1988.	160416.
PF	25-FEB-1988.	US-160416.
PR	13-JUL-1989.	US-379076.

Query Match: 9.5%; Score 183; DB 5; Length 326;
Best Local Similarity 31.5%; Pred. No. 3,26e-06;
Matches 28; Conservative 20; Mismatches 32; Indels 9; Gaps 5

Dd 40 ctscpggyasrlcgpasdvctscpknetflastnhabacvscrgtcghlssesgcdkt 99
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 28 CSNCDAGTF---CDNNRNQICSPCPNPSFSAGQ-RTDCIDCR-QCKGVFRTTRKCSST 81

Dd 100 rdrcvdcasgnycllkqgcgrtcaptk 128
:||||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 82 SNAECDCTPFFHCL--G-AGCSMCEDOCK 107

RESULT 13
ID R85071 standard; Protein; 325 AA.
AC R85071;
DT 16-APR-1996 (first entry)
DE Shope fibroma virus T2 protein.
KW Shope fibroma virus; rabbit fibroma virus; immunosuppressive;
KW cytokine antagonist; tumour necrosis factor antagonist;
KW therapeutic; cachexia; septic shock.
DB Rabbit fibroma virus.
UNUS464938-A.
PD 07-NOV-1995.
PE 18-AUG-1994; 292549.
PR 19-OCT-1992; US-963330.
PA (IMMV) Immunex Corp.
PI Goodwin RG, Smith CA
PW: 95-402861/51.
DR N-PSDB: T02470.

PT New soluble viral proteins that bind tumour necrosis factor - for treating cachexia, septic shock, side effects of TNF therapy etc., also useful in assays, affinity purification, and antibody prod.
FT Claim 1; Columns 25-28; 21pp; English.
CC The Shope fibroma virus T2 protein is a soluble viral protein which binds tumour necrosis factor (TNF), thus inhibiting TNF binding to its receptor. T2 protein is useful (when administered by injection or infusion from sustained release implants, etc.) for treating TNF associated toxicity e.g. side effects of using TNF as an antitumour agent or in cases of cachexia and septic shock where TNF production is excessive.
SO Sequence 325 AA;

Dd 49 asrlcgpsntvcspcdedgtftasnhabavscrgptghlssesgcdtrhdvncst 108
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 33 AGTFDDNNRNQICSPCPNPSFSAGQ-RTDCIDCR-QCKGVFRTTRKCSSTSNAECDCTP 90

Dd 109 gnycllkxgnagcrilcaptk 128
|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy 91 GFHCL-G-AGCSMCEDOCK 107

RESULT 14
ID R27865 standard; Protein; 325 AA.
AC R27865;
DT 11-MAR-1993 (first entry)
DE Rabbit fibroma virus T2 protein.
KW Polymerase chain reaction; TNF antagonist; tumour necrosis factor.
OS Rabbit fibroma virus.
PN WO9217583-A.
PD 15-OCT-1992.
PF 29-MAR-1991; U02207.
PR 29-MAR-1991; WO-U22076.
PA (IMMV) IMMUNEX CORP.
PI Goodwin RG, Smith CA.
PW: 92-366255/44.
DR N-PSDB: 029744.
PT Isolated viral proteins are cytokine (e.g. TNF) antagonists - for regulating immune response and for treating cachexia septic shock

```

PT and side effects associated with cytokine therapy
PS Claim 2, Page 25, 39pp; English.
CC The sequence is that of rabbit fibroma virus T2 protein which acts as
CC a TNF antagonist by binding TNF and preventing it from binding to TNF
CC receptors. It can be used for regulating immune responses, treating
CC cachexia or septic shock or to treat side effects associated with
CC cytokine therapy, e.g. TNF anti-tumour therapy. It may also be used
CC as an immunogen, a reagent in assays and as a binding agent for
CC affinity purifications. See also R2/866.
SQ Sequence 325 AA;

Query Match 9.0%; Score 174; DB 5; Length 325;
Best Local Similarity 31.3%; Pred. No. 1.75e-05;
Matches 25; Conservative 15; Mismatches 35; Indels 5; Gaps 4

Db 49 asrlcpgsnvscpcedqgtftaetnhapacsvscrgptqhlisesqpdthdivncst 108
| : | | | | | | | | | | : | | | | | | | | | | : | | :
Oy 33 AGTFCDNRNRNIGSCPSPNSFSAGGQ-RTCDICR-QCKGVFRTRKCSSRSMNECCDTP 90
| | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy 91 GFHCL-G-AGCSMCEODCK 107

RESULT 15
ID R85073 standard; Protein; 355 AA.
AC R85073;
DT 16-APR-1996 (first entry)
DE Cowpox virus T2-equivalent protein.
KW Cowpox virus; Immunosuppressive; cytokine antagonist;
KW tumour necrosis factor antagonist; therapeutic; cachexia;
KW septic shock.
OS Cowpox virus.
PN US5464938-A.
PD 07-NOV-1995.
PF 18-AUG-1994; 292549.
PR 19-OCT-1992; US-963330.
RA (IMMV ) Immunex Corp.
PI Goodwin RG, Smith CA
P1 WPI; 95-402861/51.
DR N-PSDB: T02472.

PT New soluble viral proteins that bind tumour necrosis factor - for
PT treating cachexia, septic shock, side effects of TNF therapy etc.,
PT also useful in assays, affinity purification and antibody prodn.
PS Claim 1; Colunus 35-38; 21pp; English.
CC The cowpox virus T2-equivalent protein is a soluble viral protein
CC which binds tumour necrosis factor (TNF), thus inhibiting TNF
CC binding to its receptor. T2 protein is useful (when administered
CC by injection or infusion from sustained release implants, etc.) for
CC treating TNF associated toxicity e.g. side effects of using TNF as
CC an antitumour agent or in cases of cachexia and septic shock where
CC TNF production is excessive.
SQ Sequence 355 AA;

Query Match 8.5%; Score 163; DB 15; Length 355;
Best Local Similarity 34.9%; Pred. No. 1.34e-04;
Matches 30; Conservative 21; Mismatches 24; Indels 11; Gaps 8;

Db 42 clscpgtyasrlsdskntntqgtptasctftrmhlpaciscngrcdsngyettsn 101
| : | | | | | | | | | | : | | | | | | | | | | : | | :
Oy 28 CSNCPAGF---CDNNRN-QI-CSPCPNFSFSSAGGQ-RTCDICR-QCKGVFRTRKES 79
| | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 102 tthnricdcapgyycflkgsagkac 127
| : | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 80 STSNMEDCDCTGFHC-L-GA-GCSMC 102

Search completed: Tue Mar 24 17:01:54 1998
Job time : 98 secs.

```

MPERCH

(TM)

Release 2.1D John F. Collins, BioComputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelligentGenetics, Inc.

MPerch_n n.a. - n.a. database search, using Smith-Waterman algorithm
n: Wed Mar 25 13:33:00 1998; MasPar time 697.53 Seconds
1001.767 Million cell updates/sec
Tabular output not generated.

Title: >US-08-236-918A-7
(1-1415) from 5674704.seq
Description: 1415
Perfect Score: 1415
N.A. Sequence: 1 AGTGAAGACTCTCCGGCAG.....TGATAAAAA.....AAAAA 1415
Comp: TCACCTTCAGAGAGCCGTC.....ACATATTTT.....TTTTTTT

Scoring table: TABLE default
Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 665703 seqs, 246912890 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

EST-A
1: EST1 2: EST2 3: EST3 4: EST4 5: EST5 6: EST6 7: EST7 8: EST8
9: EST9 10: EST10 11: EST11 12: EST12 13: EST13 14: EST14
15: EST15 16: EST16 17: EST17 18: EST18 19: EST19 20: EST20
21: EST21 22: EST22 23: EST23 24: EST24 25: EST25 26: EST26
27: EST27 28: EST28 29: EST29 30: EST30 31: EST31 32: EST32
33: EST33 34: EST34 35: EST35 36: EST36 37: EST37 38: EST38
39: EST39 40: EST40 41: EST41 42: EST42 43: EST43 44: EST44
45: EST45 46: EST46 47: EST47 48: EST48 49: EST49 50: EST50
51: EST51 52: EST52 53: EST53 54: EST54 55: EST55 56: EST56
57: EST57 58: EST58 59: EST59 60: EST60 61: EST61 62: EST62
63: EST63 64: EST64 65: EST65 66: EST66 67: EST67 68: EST68
69: EST69 70: EST70 71: EST71 72: EST72 73: EST73 74: EST74
75: EST75 76: EST76 77: EST77 78: EST78 79: EST79 80: EST80
81: EST81 82: EST82 83: EST83 84: EST84 85: EST85 86: EST86
87: EST87 88: EST88 89: EST89 90: EST90 91: EST91 92: EST92
93: EST93 94: EST94 95: EST95 96: EST96 97: EST97 98: EST98
99: EST99 100: EST100 101: EST101 102: EST102 103: EST103
104: EST104 105: EST105 106: EST106 107: EST107 108: EST108
109: EST109 110: EST110 111: EST111 112: EST112 113: EST113
114: EST114 115: EST115 116: EST116 117: EST117 118: EST118
119: EST119 120: EST120 121: EST121 122: EST122 123: EST123
124: EST124 125: EST125 126: EST126 127: EST127 128: EST128
129: EST129 130: EST130 131: EST131 132: EST132 133: EST133
134: EST134 135: EST135 136: EST136 137: EST137 138: EST138
139: EST139 140: EST140 141: EST141 142: EST142 143: EST143
144: EST144 145: EST145 146: EST146 147: EST147 148: EST148
149: EST149 150: EST150 151: EST151 152: EST152 153: EST153
154: EST154 155: EST155 156: EST156 157: EST157 158: EST158
159: EST159 160: EST160 161: EST161 162: EST162 163: EST163
164: EST164 165: EST165 166: EST166 167: EST167 168: EST168
169: EST169 170: EST170 171: EST171 172: EST172 173: EST173
174: EST174 175: EST175 176: EST176 177: EST177 178: EST178

Database:

EST-B
99: EST99 100: EST100 101: EST101 102: EST102 103: EST103
104: EST104 105: EST105 106: EST106 107: EST107 108: EST108
109: EST109 110: EST110 111: EST111 112: EST112 113: EST113
114: EST114 115: EST115 116: EST116 117: EST117 118: EST118
119: EST119 120: EST120 121: EST121 122: EST122 123: EST123
124: EST124 125: EST125 126: EST126 127: EST127 128: EST128
129: EST129 130: EST130 131: EST131 132: EST132 133: EST133
134: EST134 135: EST135 136: EST136 137: EST137 138: EST138
139: EST139 140: EST140 141: EST141 142: EST142 143: EST143
144: EST144 145: EST145 146: EST146 147: EST147 148: EST148
149: EST149 150: EST150 151: EST151 152: EST152 153: EST153
154: EST154 155: EST155 156: EST156 157: EST157 158: EST158
159: EST159 160: EST160 161: EST161 162: EST162 163: EST163
164: EST164 165: EST165 166: EST166 167: EST167 168: EST168
169: EST169 170: EST170 171: EST171 172: EST172 173: EST173
174: EST174 175: EST175 176: EST176 177: EST177 178: EST178

179: EST179 180: EST180 181: EST181 182: EST182 183: EST183
184: EST184 185: EST185 186: EST186 187: EST187 188: EST188
189: EST189 190: EST190 191: EST191 192: EST192 193: EST193
194: EST194 195: EST195 196: EST196

Statistics: Mean 11.971; Variance 4.024; scale 2.975

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	174	12.3	610	191	AA109726	mp10f11.r1 Life Tech	7.20e-165
2	153	10.8	589	166	AA087107	mol3601.r1 Life Tech	1.15e-139
3	127	9.0	278	14	R06117	y889b09.s1 Homo sapie	7.32e-109
4	120	8.5	388	14	R06116	y889b09.s1 Homo sapie	1.15e-100
5	119	8.4	330	116	HSDH25C04	H. sapiens partial cd	1.68e-99
6	114	8.1	408	77	N32944	y10b06.s1 Homo sapie	1.11e-93
7	114	8.1	424	146	AA13170	zm27e06.s1 Stratagene	1.11e-93
8	115	8.1	571	148	AA131088	zol6c04.s1 Stratagene	7.62e-95
9	113	8.0	353	26	R44830	y31d01.s1 Homo sapie	1.60e-92
10	113	8.0	471	63	H24051	ym34c03.s1 Homo sapie	1.60e-92
11	113	8.0	609	178	AA194481	zq05d10.s1 Stratagene	1.60e-92
12	111	7.8	426	97	N80739	z98c06.s1 Homo sapie	3.31e-90
13	111	7.8	461	185	N73439	y231g10.r1 Soares mul	3.31e-90
14	109	7.7	300	6	T72377	yc1e09.s1 Homo sapie	6.79e-88
15	109	7.7	438	144	AA011026	ze34c08.s1 Soares ret	6.79e-88
16	108	7.6	386	174	AA172046	zo96c01.s1 Stratagene	9.68e-87
17	108	7.6	433	177	AA188423	zp78b07.s1 Stratagene	9.68e-87
18	108	7.6	566	147	AA126763	zn87b11.s1 Stratagene	9.68e-87
19	106	7.5	398	132	N95817	zb66c08.s1 Soares fet	1.95e-84
20	106	7.5	617	177	AA191659	zq43b02.s1 Stratagene	2.76e-83
21	105	7.4	369	177	AA190727	zp88b04.r1 Stratagene	2.76e-83
22	105	7.4	396	44	R25141	yh36g09.r1 Homo sapie	2.76e-83
23	105	7.4	440	87	N66948	z488e10.r1 Homo sapie	2.76e-83
24	104	7.3	370	50	M77984	ES01568 Homo sapiens	3.88e-82
25	104	7.3	338	164	AA079507	zm97407.s1 Stratagene	3.88e-82
26	104	7.3	381	65	H30221	yp41h12.r1 Homo sapie	3.88e-82
27	104	7.3	410	67	H38146	yp37a10.s1 Homo sapie	3.88e-82
28	104	7.3	424	21	H01809	zo20h09.s1 Stratagene	3.88e-82
29	104	7.3	432	149	AA134241	y132h07.r1 Homo sapie	3.88e-82
30	104	7.3	424	21	AA205931	zq53a05.s1 Stratagene	3.88e-82
31	104	7.3	450	81	H62775	yt66g07.r1 Homo sapie	3.88e-82
32	103	7.3	496	183	N53968	yy60e11.s1 Soares fet	5.45e-81
33	103	7.3	542	164	AA081138	zn34h07.s1 Stratagene	3.88e-82
34	103	7.3	563	182	AA209208	zq63e09.s1 Stratagene	3.88e-82
35	103	7.3	586	147	AA126811	zn87b11.r1 Stratagene	5.45e-81
36	102	7.2	371	16	T08745	ES00637 Homo sapiens	7.63e-80
37	102	7.2	379	174	AA19163	zp12a02.s1 Stratagene	7.63e-80
38	102	7.2	412	168	AA12432	zn24g01.s1 Stratagene	7.63e-80
39	102	7.2	440	99	N94134	za53h01.r1 Homo sapie	7.63e-80
40	102	7.2	442	55	R97235	yq92g11.r1 Homo sapie	7.63e-80
41	101	7.1	325	15	T05070	ES002958 Homo sapiens	1.06e-78
42	101	7.1	398	91	H75272	yn07g01.r1 Homo sapie	1.06e-78
43	101	7.1	410	25	R38419	yf88b09.s1 Homo sapie	1.06e-78
44	101	7.1	463	99	N94342	zb57f06.s1 Homo sapie	1.06e-78
45	101	7.1	522	9	T86817	yt66h07.r1 Homo sapie	1.06e-78

ALIGNMENTS

RESULT 1
LOCUS AA109726 610 bp mRNA
DEFINITION mp10f11.r1 Life Tech mouse embryo 8 5gpc 10664019 Mus musculus cDNA
clone 568845 5' similar to gb:J04492 Mouse T-cell receptor 4-1BB
ACCESSION AA109726
NID AA109726
KEYWORDS EST.

[illegible]

QY	529	CAAACTGTTCTTGGATGCAGGAAGTCGTCTTGCTTGATGGACGACGAGCAGCGTGG	588
Db	552	tgtgttgacc 561 	
QY	589	TCTGTGACACC 598	
RESULT	2		
LOCUS	AA087107	589 bp	mRNA EST 23-Oct-1996
DEFINITION	mol3g91.r1 Life Tech mouse embryo 10 5dpc 10665016 Mus musculus cDNA clone 553488 5' similar to gb:304492 Mouse T-cell receptor 4-1Bb protein mRNA, complete cds (MOUSE)..		
ACCESSION	AA087107		
NID	g1630565		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
REFERENCE	1 (bases 1 to 589)		
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubnue,T., Geller,S., Kucher,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Weisling,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.		
TITLE	The WashU-HMNI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	Contact: Marra M/Mouse EST Project WashU-HMNI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mousestewatson.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MG1:334280 Seq primer: -28M13 rev1 from Amer sham High quality sequence stop: 377. Location/Qualifiers 1..589 /organism="Mus musculus" /strain="C57BL/6J" /note="Vector: PCMV-SPORT2; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. 10.5dpc embryos. PCMV-SPORT2 vector." /clone="553488" /clone_lib="Life Tech mouse embryo 10 5dpc 10665016" /dev_stage="10.5dpc embryos" /lab_host="DH10B" <1..>589		
FEATURES			
SOURCE			
BASE COUNT	133 a	138 c	171 g 147 t
ORIGIN			
	mRNA		
Query Match	10.8%; Score 153; DB 166; Length 589;		
Best Local Similarity	69.2%; Pred. NO.1.15e-139;		
Matches 339; Conservative	0; Mismatches 144; Indels 7; Gaps 7;		
Db	76 attcgccatcgggaacaactgtttacaaagtgtgtgatctgtgtcgtagtgagtc 135 		
QY	112 ATTTCATCATGTGGAAGAAGCGTGTACAACATAGTAGCACTCTGTGCTGCTCAACT 171 		
Db	136 gtgagaagvtgvgagccgcgtgcagaactccctgtgtataactctcagcctgttacttctgc- 194 		
QY	172 TTGAGAGGACGAAGATCATTTGCAGAGATCCTTGTAATACTGCCACGCTGGTACATTCTGTG 231 		
Db	195 agaa-aatacaatccag-tctgcaagaagctgcacctccaagaactcttcasagcatagtg 252 		
QY	232 ATAAATACAGGAATCAGATTGGCAGTCCCTGTCTCCCAAAATAGTTCTCCAGCGCAGGTG 291 		

Db	253	gaagcgaactgtaacatcttcacagafgltgagcggcatttcaggttc	312
QY	292	GACCAAGAACCTGTGTACATATGCACGACGTGTAAGTGTTCACGAGCACAGAAAGAGT	351
Db	313	gtccctctaccacaacacgaggtgtgagtcgattgaaagatccatcgcttggtggccac	372
QY	352	GTTCCTCCACCAAGCATATGCAGAGTGTGACTGCACTCCAGGGTTTACGTCCCTGGGGGCGAG	411
Db	373	atgcaaccagatgtgaaagagactcgaagcctcgaagaaactcaagaaagcaggttgc	432
QY	412	GATGACAGCATGTGTGAACAGAGTTGTAACAAGTCAAGAA-CTGCAAAAAAAGGTGT	470
Db	433	aaacactgtaactctgggacacatttaatgcacgaacggaaggtgtgtctgcacccctgg	492
QY	471	AAAGCATGTTGCTTTGGGACATTTTAAACATCAGAAACGT-GGCAT-CTGTCAACCTGGA	528
Db	493	cgaggtgcgtctatcttatagtcttgcctaa-gaccgggacacggagaagagacgtgg	551
Db	529	CAAACTGTTCTTTGGAGTGAAGTGTGTGCTTGTAATGGACGAAGGAGAAGGACGTGG	588
QY	589	TCTGTGGACC 598	

RESULT	3				
LOCUS	R06117	278 bp	mRNA	EST	03-APR-1995
DEFINITION	ye9b09.s1 Homo sapiens cDNA clone 124889 3'				similar to contains
	Alu repetitive element.				

ACCESSION	R06117
NID	9756737
KEYWORDS	EST.

SOURCE human clone-124889 library-Souares fetal liver spleen hNF1S
human pPRT3D (Pharmacia) with a modified polylinker hNF3-
(ampicillin resistant) primer-2mls1 RstlelPac I Rstle2-Eco RI
Liver and spleen from a 20 week-post conception male fetus. 1st
strand cDNA was primed with a Pac I - oligo(dT) primer [5],
ACTCGACAGATTAATTAAGAATCTTTTTTTTTTTTTTTT 3', double-stranded
cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
I and cloned into the Pac I and Eco RI sites of the modified pRT3
vector. Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonalolo.

ORGANISM
Rufo sapiens
 Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
1 (bases 1 to 278)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, C., Marra, M.

Parsons, J., Rifkin, J., Rofling, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and Waterston, R., 1996. *Genome*. (New York: Cold Spring Harbor Press).

TITLE	Wilson, R.
JOURNAL	The Washu-Merck EST Project Unpublished (1995)

COMMENT

Contact: Wilson RK
w@r.com

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watscn.wustl.edu

High quality sequence stops: 257
Source: IMAGE Consortium, LNL

FEATURES
This clone is available royalty-free through LINL ; contact the
IMAGE Consortium (info@image.linl.gov) for further information.
Location/Qualifiers

```
source  
1.:278  
/organism="Homo sapiens"  
61000-#124800
```

BASE COUNT	/C10NE- 124009			
ORIGIN	85 a	60 c	76 g	56 t
				1 others

Query Match 9.08; Score 127; DB 14; Length 278;

[illegible]

		RESULT	4	P0617C	300 LBS	—PWT	TOTL	80	90	100	110	120
--	--	--------	---	--------	---------	------	------	----	----	-----	-----	-----

LOCUS	388 bp	mRNA	03-APR-1995
DEFINITION	Y88b09.r1 Homo sapiens CDNA clone 124889 5' similar to contains Alu repetitive element;.		

ACCESSION	R06116
NID	9756736
KEYWORDS	EST

ALTERNATES	SOURCE
EST1	human
	clone=124889 library=Soares fetal liver spleen INFIL
	vector=pt713D (Pharmacia) with a modified polylinker host=DH10B

(ampici
Liver a
strand

AACTGGAGAGATTAATTAAGATCTTTTCTTTTCTTTT 3'), double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pRT3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fátima Bonalfo.

ORGANISM Homo sapiens

Eucaryotes: Metazoa: Chordata: Vertebrata: Gnathostomata: Mammalia

Eutheria: Primates: Catarrhini: Hominoidea: Homo.

(bases 1 to 388)

REFERENCE Hillier, L., Clark, N., Dubouque, T., Elliston, K., Hawkins, M.,

Holman, M., Hultman, M., Kuucaba, T., Le, M., Lennon, G., Marita, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and

WILSON, R.
The Washu-Merck EST Project
Newbl: 1963 (1005)

COGNOMIAL	unpublished (1993)
COMMENT	Contact: Wilson R.

WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
High quality sequence stops: 312
Source: IMAGE Consortium, LINT

This clone is available royalty-free through LINL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

```

      1.388
      /organism="Homo sapiens"
      location/qualifiers
      source
      features

```

BASE COUNT	88 a	92 c	85 g	120 t	3 others
ORIGIN	./clone="124889"				

Query Match	Score	DB	Length
8.58;	120;	14;	388;

Best Local Similarity	/5.8*	Pred. NO. 1.15e-100;			
Matches	188; Conservative	0; Mismatches	58; Indels	2; Gaps	2

```
Db      134 ttttttttttgaagcagggtctcaactctgtcaaccaggctnagtacatgattgtg 193
```

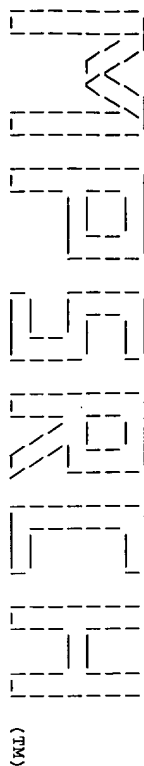
Oy	1059	TTTTTTT	TTTTTTT	TGACAGGGGTCTACTCTGTGCACCCAGGCGGAGGTGCAATGGGACC	1118
Db	194	atcatglttcacgcagcttcgcacctctgg-gttcaagtatcctccaccatagcctc	252		
Oy	1119	ACCAAGGCTCTCTGCAGCCTTGACTTCGGAGGCTCAAGTATCATCTCTGGCTAGGCTC	1178		
Db	253	cgaactgcttgagctaacaggcgacgcattcacacaanccagctaatttttgtatttttt	312		
Oy	1179	CTTAGTGTGGAACTACACAGAGAGGCCACACACCTGACTACTTATTTTTGT	1238		
Db	313	tgtaaacctgggggtgttgccatgttgcaccaagggtgtttccaacttcctaagtc	372		
Oy	1239	TGCTAAGAATGGCANTTGGCCATGTTGTACAGAGGCGGTCT-CAACATCTAGTTCATT	1297		
Db	373	aatcctcc 380			
Oy	1298	TGGCTCC 1305			
RESULT	5	HSDH25C04	330 bp	RNA	EST
LOCUS		H. sapiens partial cDNA sequence; clone H25C04; single read.			
DEFINITION		221713			
KEYWORDS		938560			
SOURCE		partial cDNA sequence; transcribed sequence fragment.			
ORGANISM		human.			
REFERENCE		Homo sapiens			
AUTHORS		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;			
TITLE		Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
JOURNAL		1 (bases 1 to 307)			
REFERENCE		Genexpress.			
AUTHORS		Direct Submission			
TITLE		Submitted (22-FEB-1993) Genzentrum Muenchen, Laboratorium fuer			
JOURNAL		molekulaere Biologie, Am Klopferplatz 18a, 8033 Martinsried,			
COMMENT		Germany. E-mail: obermaier@vms.biochem.mpg.de			
FEATURES		2 (bases 1 to 330)			
SOURCE		Genexpress.			
LOCATION/Qualifiers		The Genexpress cDNA program			
1..330		Unpublished			
/organism="Homo sapiens"		cloning vector is pBluescript SK(+);			
/clone_id="Stratagene cDNA library Human heart,		Genexpress library reference is A.			
cat#936208"		automatic.			
BASE COUNT		109 a 66 c 96 g 59 t			
ORIGIN					
Very Match		8.4%; Score 119; DB 116; Length 330;			
Match Local Similarity 78.2%; Pred. No. 1.68e-99;					
Matches 186; Conservative 0; Mismatches 49; Indels 3; Gaps 3;					
Db	12	tgaagtcagaggttggaaaccagctgtgccacatgtgnaaaaaccctactactaaaa	71		
Cp	1294	TGAACCTTAGGAGTTTGAACAACGCTGTACAACTAGGGAATGCCATTTTACCAACA	1235		
Db	72	agaataatacaaaatltagccaagtgtgtgtgcgcttacctttaaccacagctagtgyga	131		
Cp	1234	A-AAAAACAAAAAAGTATATCATCAGTGTGTGGCCCTCTCTTAAGTTCACAGCTACTAGG-A	1177		
Db	132	gactyagcaggaagaagcacttgaac-cttyagagcgaggttgcagtlyagcagaatca	190		
Cp	1176	GACTGAGGACAGAGAGATCATTTGAGCTCCAGAGAGGTCAAGGCTCACAGACCCATGCTGG	1117		
Db	191	agccaactgcatccagccttgggtgtgacagaagaagatccactctggaaaaaaaaaaaaa	248		
Cp	1116	TGCCACTGCTACTCCAGCCTGGGTGTGACAGATGAGACCTGTCAAAAAAAAAAAAAA	1059		
RESULT	6	N32944	408 bp	mRNA	EST
LOCUS					
					10-JAN-1996

DEFINITION	1710b06.s1 Homo sapiens cDNA clone 270803 3' similar to contains Alu repetitive element..					
ACCESSION	N32944					
NID	g1153343					
KEYWORDS	EST.					
SOURCE	human clone-270803 primer-m13 -40 forward library=Soares melanocyte 2NHM vector-pRTT3D (Pharmacia) with a modified polylinker host-PHI0B (ampicillin resistant) RstE1-Not I RstE2-Eco RI Mael 1st strand cDNA was primed with a Not I - O190(dT) primer [5'-TGTTCCACTCTGAAAGTAGGGAGCGGCCGCAGTTTTTTTTTTTTTTT-3'] double-stranded cDNA was size selected, ligated to Eco RI adapted (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pRTT3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonardo. RNA from normal forestin melanocytes (FS3/4) was kindly provided by Dr. Anthony J. Albino.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryotes; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 408)					
AUTHORS	Hallier,L., Clark,N., Dubucq,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kuebb,T., Le'M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasaki,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.					
TITLE	The WashU-Merck EST Project					
JOURNAL	Unpublished (1995)					
COMMENT	Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu High quality sequence stops: 394 Source: IMAGE Consortium, LNU, This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Location/Qualifiers 1. 408					
FEATURES	source /organism="Homo sapiens" /cclone="270803"					
BASE COUNT	mRNA 93 a 101 c 96 g 118 t					
ORIGIN						
Query Match	8.1%; Score 114; DB 77; Length 408;					
Best Local Similarity	77.4%; Pred. No. 1,1le-93;					
Matches	182; Conservative 0; Mismatches 50; Indels 3; Gaps 3					
Db	4 tttttttttttaaagcaggctcacctgtaccttgtaactaaagtcggggtgaagcccaa 63					
Oy	1059 TTTTtt					
Db	64 atcatggtcactcgaagcttcaattccctcg-gctcaagtatcttcctgcgtcaagcct 122					
Oy	1119 ACCAAGGCTCtGTGAGCGCTTCAGCTCGGAGGTCAAGTATGATCCTCGCTCAGTCT- 1177					
Db	123 ccagtaggtggagctacagaagaagaccacacacccccagctcgatttttatattttc 182					
Oy	1178 CCTACTACTCTGAAGTAACAAGAGGGGCGCACACACCTGACTTAATTtttttttttTGT 1237					
Db	183 ttg-taaagaagaagctctgctaagtatgaagctattaattgaactcctgggt 236					
Oy	1238 TTGGTAAAGAAGGCGCATTTTCGCCCAAGTGTGTACAGGCTGCTCCAAGTCTTAGGTT 1292					
RESULT	7					
LOCUS	AA113170 424 bp mRNA EST 08-NOV-1996					
DEFINITION	zm27e06.sl Stratiogene pancreas (#937208) Homo sapiens cDNA clone					

ACCESSION	526882.37	similar to contains Alu repetitive element.
NID	AA113170	
KEYWORDS	g1664/73	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 424)	
AUTHORS	Hallier, L., Clark, N., Dubouque, T., Elliston, R., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevisakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.	
TITLE	WashU-Merck EST project	
JOURNAL	Unpublished (1995)	
COMMENT	Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu This clone is available royalty-free through LHM; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: -40M13 fwd. from Arabidopsis High quality sequence stop: 386. Location/Qualifiers 1..424	
FEATURES	/organism="Homo sapiens" /note="Organ: pancreas; Vector: pBluescript SK-; Site_1: ECORI; Site_2: XhoI; Cloned unidirectionally. Primer: Inset size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCAGGAG 3' -3' adaptor sequence: 5' CTCGACTTTTCTTTTTTTTTTTT 3'" /clone_id="Stratagene pancreas (#937208)" /lab_host="SOLR cells (kanamycin resistant)" complement(<1..>424)	
BASE COUNT	96 a 110 c 81 g 135 t 2 others	
ORIGIN	MRNA	
Query Match	Best Local Similarity 78.3%; Score 114; DB 146; Length 424;	
Matches	184; Conservative 0; Mismatches 47; Indels 4; Gaps 4;	
Dy	tcttttttttttgcagaagggtcactctgttaacctagctgggtgcagtgcna 63	
Oy	TTTTTTTTTTTTTTTTTGACAGGGGTCTCACTCTGCACCAGCGTAGTGCAGGCC 118	
Dd	atcatggtcactcagcctgaattcccttg-gtcaagtagtatcttcctgaacct 122	
Oy	1119 ACCAGGCTCTCTCGAGCCCTTGACCTCGGAGGCTCAAGTAGATCTCTCGCTCAGTCT- 1177	
Dd	cacagtagtctggaactacagaagaaccaccacacaccagctga-tttttatcttct 181	
Oy	1178 CCGTAGTACTGGACTACAAAGAAGGCGCACACACCTGATTAATTCTTTTGTGTTTGTG 1237	
Dd	ttg-taaagacaagagctctgactatggtatgatcagaagctattttaactcctgggt 235	
Oy	1238 TTGTGAAGAAGAGCATTTTCGCCAATGTTGTACAGAGCTGTGTCACAACCTCAGAGTT 1292	
RESULT	8	
LOCUS	AA131088	571 bp mRNA EST 27-NOV-1996
DEFINITION	zo16C04.31 Stratagene colon (#937204) Homo sapiens CDNA clone	
ACCESSION	AA131088	
NID	g1692578	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

[illegible]

THIS PAGE BLANK (USPTO)



Release 2.1D John F. Collins, BioComputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Merch_nam n.a. - n.a. database search, using Smith-Waterman algorithm

Run: Wed Mar 25 13:46:34 1998; Maspar time 378.76 Seconds
1053.585 Million cell updates/sec

Output not generated.

Title: >US-08-236-918A-7
Description: (1-1415) from 5674704.seq
Perfect Score: 1415
N.A. Sequence: 1 AGCGAAGTCTCCGCGAC.....TGATAAAAAAAAAAAAAA 1415
Comp: TCACCTTCAAGAGGCCGCTC.....ACTATTTTTTTTTTTTTTTT

Scoring table: TABLE default
Gap 6

Match STD : Dbase 0; Query 0

Searched: 397346 seqs, 141010104 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

EST-C
1:EST197 2:EST198 3:EST199 4:EST200 5:EST201 6:EST202
7:EST203 8:EST204 9:EST205 10:EST206 11:EST207 12:EST208
13:EST209 14:EST210 15:EST211 16:EST212 17:EST213
18:EST214 19:EST215 20:EST216 21:EST217 22:EST218
23:EST219 24:EST220 25:EST221 26:EST222 27:EST223
28:EST224 29:EST225 30:EST226 31:EST227 32:EST228
33:EST229 34:EST230 35:EST231 36:EST232 37:EST233
38:EST234 39:EST235 40:EST236 41:EST237 42:EST238
43:EST239 44:EST240 45:EST241 46:EST242 47:EST243
48:EST244 49:EST245 50:EST246 51:EST247 52:EST248
53:EST249 54:EST250 55:EST251 56:EST252 57:EST253
58:EST254 59:EST255 60:EST256 61:EST257 62:EST258
63:EST259 64:EST260 65:EST261 66:EST262 67:EST263
68:EST264 69:EST265 70:EST266 71:EST267 72:EST268
73:EST269 74:EST270 75:EST271 76:EST272 77:EST273
78:EST274 79:EST275 80:EST276 81:EST277 82:EST278
83:EST279 84:EST280 85:EST281 86:EST282 87:EST283
88:EST284 89:EST285 90:EST286 91:EST287 92:EST288
93:EST289 94:EST290 95:EST291 96:EST292 97:EST293
98:EST294

Database:

EST-D
99:EST295 100:EST296 101:EST297 102:EST298 103:EST299
104:EST300 105:EST301 106:EST302 107:EST303 108:EST304
109:EST305 110:EST306 111:EST307 112:EST308 113:EST309
114:EST310 115:EST311 116:EST312 117:EST313 118:EST314
119:EST315

Statistics: Mean 11.830; Variance 3.250; scale 3.640

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
C 1	114	8.1	503	98	AA282360	zr12a11.r1 Soares MbH	3.97e-116
C 2	108	7.6	251	41	AA338322	EST43242 Fetal Brain	1.26e-107
C 3	108	7.6	300	44	AA262532	zsr1c02.r1 Soares MbH	1.26e-107
C 4	108	7.6	447	94	AA127499	zn81e05.s1 StrataGene	1.26e-107
C 5	106	7.5	514	91	C06339	similar to none.	8.32e-105
C 6	105	7.4	465	71	AA226159	nc10b07.r1 NCI CGAP	2.13e-103
C 7	104	7.3	361	72	AA229986	nc37d05.s1 NCI CGAP	5.43e-102
C 8	103	7.3	367	36	AA281211	EST29525 Cerebellum	1.38e-100
C 9	104	7.3	384	69	W93002	zn80d12.s1 Soares Fet	5.43e-102
C 10	104	7.3	426	62	AA386111	EST07525 Brain IV Hom	5.43e-102
C 11	103	7.3	428	94	AA128930	nc08d10.s1 StrataGene	1.38e-100
C 12	102	7.2	428	71	AA225537	nc08d07.s1 NCI CGAP	3.49e-99
C 13	102	7.2	468	10	AA234890	zr17h10.r1 Soares MbH	3.49e-99
C 14	100	7.1	351	93	C15363	Human Fetal Brain CDN	2.21e-96
C 15	100	7.1	379	2	AA195351	zn13g12.s1 Soares MbH	2.21e-96
C 16	100	7.1	396	70	AA224851	nc33d06.s1 NCI CGAP	2.21e-96
C 17	101	7.1	396	70	AA258786	zsr2b04.r1 Soares MbH	8.80e-98
C 18	101	7.1	422	43	AA258889	zsr2b04.r1 Soares MbH	8.80e-98
C 19	100	7.1	423	15	T03583	1B521 Infant Brain	2.21e-96
C 20	101	7.1	423	15	AA225948	nc11f01.s1 NCI CGAP	8.80e-98
C 21	101	7.1	427	80	AA404276	zvr63e03.s1 Soares tot	8.80e-98
C 22	101	7.1	498	7	AA223291	zr08d02.s1 StrataGene	2.21e-96
C 23	100	7.0	510	91	C06329	similar to none.	5.52e-95
C 24	99	7.0	303	74	AA347843	EST54212 Fetal heart	5.52e-95
C 25	99	7.0	383	37	AA329122	EST32968 Embryo, 12 w	5.52e-95
C 26	99	7.0	465	8	AA227713	zn156d09.r1 Soares MbH	5.52e-95
C 27	97	6.9	299	71	AA225709	nc10b12.r1 NCI CGAP	3.41e-92
C 28	98	6.9	309	70	AA216399	nc10f03.s1 NCI CGAP	3.41e-92
C 29	97	6.9	318	51	AA355083	EST63493 Jurkat T-cell	1.37e-93
C 30	97	6.9	362	72	AA230025	nc38a10.s1 NCI CGAP	3.41e-92
C 31	98	6.9	378	70	AA224889	nc34d11.s1 NCI CGAP	1.37e-93
C 32	98	6.9	384	52	AA358717	EST67628 Fetal lung I	1.37e-93
C 33	98	6.9	417	7	AA224323	zr12b02.s1 StrataGene	1.37e-93
C 34	96	6.8	309	70	AA216345	nc10b12.s1 NCI CGAP	8.40e-91
C 35	96	6.8	318	71	AA225950	nc11f04.s1 NCI CGAP	8.40e-91
C 36	96	6.8	432	71	AA225956	nc18g12.s1 NCI CGAP	8.40e-91
C 37	95	6.7	221	41	AA338517	EST43490 Fetal brain	2.06e-89
C 38	95	6.7	272	99	AA281516	zs8d10.s1 Soares MbH	2.06e-89
C 39	95	6.7	288	83	AA420745	nc33d03.s1 NCI CGAP	2.06e-89
C 40	95	6.7	300	83	AA420806	nc33d03.s1 NCI CGAP	2.06e-89
C 41	95	6.7	330	71	AA225173	nc11g08.s1 NCI CGAP	2.06e-89
C 42	95	6.7	427	88	AA102241	zn10g09.r1 StrataGene	2.06e-89
C 43	95	6.7	452	71	AA225230	nc32b01.r1 NCI CGAP	2.06e-89
C 44	95	6.7	457	98	AA284247	zc65d01.r1 Soares Fet	2.06e-89
C 45	94	6.6	291	27	AA302682	EST11062 Umbilical ve	5.04e-88

ALIGNMENTS

RESULT 1
LOCUS AA282360 503 bp mRNA
DEFINITION zrl2a11.r1 Soares NdbHGC Homo sapiens cDNA clone 712917 5' similar to contains Alu repetitive element, contains element PTR7 repetitive element ;.

ACCESSION AA282360
NID G1925276
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryotic; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 503)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.
TITLE WashU-Merck EST Project 1997

KEYWORDS EST.
human.

SOURCE Homo sapiens
ORGANISM Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 300)

REFERENCE Hillier, L., Clark, N., Dubnue, T., Ellison, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Riklin, L., Rohlfing, T., Soares, M., Tan, F., Tevaskis, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.

AUTHORS The WashU-Merck EST Project
Unpublished (1995)

TITLE Contact: Wilson RK
JOURNAL WashU-Merck EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LIND ; contact the IMAGE Consortium (infoimage.llnl.gov) for further information.
Putative full length read
The vector to vector length is
Seq primer: -28ml3 revE ET from Amersham.
Location/Qualifiers
1..300
/organism="Homo sapiens"
/note="Vector: pT7TD-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from germinal B-cells (flow-sorted from tonsils) provided by Dr. Louis Staudt of the NCI, and was then primed with a Not I - oligo(dT) primer [5' TGTTACCAATCGAAGTGGAAGGCGCCGCCTATTATTTTTTTTTTTT 3'] . Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7D3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
/clone="685442"
/clone_lib="Soares NBHTBC"
/rissue_type="Germinal B-cell"
/lab_host="DH10B"

FEATURES mRNA
BASE COUNT 86 a 71 c 77 g 66 t
ORIGIN
<1...>300

Match 7.6%; Score 108; DB 44; Length 300;
Local Similarity 76.2%; Ped. No. 1.26e-107;
Matches 179; Conservative 0; Mismatches 53; Indels 3; Gaps 3;

Db	68	tggagagcctgcttagcccggaatttgagagcccaactgsgccaagtatgagacccta	127
Cp	1307	TGGAGGCCAAAGTAACCTAGGAGATTGAGACCACCTGTACAACATGGCGAATGCCA	1248
Db	128	tctcatatcaa-aaaatatcaaaaaaataggcagaagcgtaattgaggcacacctgtaatcc	186
Cp	1247	TCTTTACCAAAACAAAAAACAAAAGTTTAGTACAGTGTCGTGGCCCCCTCCTTGATGTTCC	1188
Db	187	agctactcagaagcgtgsgcgaagggttcacttgagc-ccagsgatccagagctgtaat	245
Cp	1187	AGCTACT-AGGAGAAGTGAAGCAGCGAGATCACTTGAGCTCCACAGAGTCACAAAGTCGACGA	1129
Db	246	gagcatgatattgggcgaactgacacctagcctctgggagatagatgtagaccctgtca	300
Cp	1128	GAGCATGATGTCGCCACTGCACCTCACCGCTGGGTGACAGAGATGAGACCTGTCA	1074

RESULT 4 AA127499 447 bp mRNA EST 19-MAY-1997
LOCUS ZN18E05.31 Striatogene lung carcinoma 937218 Homo sapiens cDNA clone
DEFINITION 564608 s' similar to contains Alu repetitive element; contains

ACCESSION	element MER22 repetitive element ;					
NID	A127499					
KEYWORDS	g1686789					
SOURCE	EST.					
ORGANISM	human.					
	Homo sapiens					
	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;					
	Vertebrata; Mammalia; Eutheria; Primates; Carnivora; Homnidae;					
REFERENCE						
AUTHORS	Homo. 1 (bases 1 to 447) Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,B., Rohlfing,T., Tan,F., Trevisakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R. Washu-Merck EST Project Unpublished (1995)					
TITLE						
JOURNAL						
COMMENT						
	Contact: Wilson RK					
	Washu-Merck EST project					
	Washington University School of Medicine					
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108					
	Tel.: 314 286 1800					
	Fax: 314 286 1810					
	Email: est@watson.wustl.edu					
	This clone is available royalty-free through LBNL ; contact the					
	IMAGE Consortium (info@image.lbl.gov) for further information.					
	Insert Length: 511 Std Error: 0.00					
	Seq primer: -40M13 fwd. from Amersham					
	High quality sequence stop: 409.					
FEATURES						
SOURCE						
	Location/Qualifiers					
	1..447					
	/organism="Homo sapiens"					
	/note="vector: pBluescript SK-; Site_1: EcoRI; site_2:					
	XhoI; Cloned unidirectionally." Primer: Oligo dT. Small					
	cell carcinoma cell line NCI-H69. Average insert size:					
	1.0 kb. Uni-ZAP XR Vector: -5' adaptor sequence: 5'					
	GAAATTCGCACGAG 3'-3' adaptor sequence: 5'					
	CCTCAGATTCTTTTCTTTTTTTT 3'"					
	/clone="564608"					
	/clone_lib="Stratagene lung carcinoma 937218"					
	/dev_stage="cell line NCI-H69"					
	/lab_host="SOLR (kanamycin resistant)"					
	/complement(4..2447)					
	/db_xref="GDB:4593662"					
BASE COUNT	97 a 122 c 89 g 139 t					
ORIGIN						
	Query Match 7.6%; Score 108; DB 94; Length 447;					
	Best Local Similarity 75.9%; Pred.No.1.2e-107;					
	Matches 180; Conservative 0; Mismatches 54; Indels 3; Gaps 3.					
Db	83	tttttttttttggagatagaagtctacactgtgcaccaggcgtggagttaattgacct	142			
Oy	1059	TTTTTTTTTTTTTTTTTGACAAGGGGTCTACTCTGTGCACCAGCGTAGTGAGGGCAAC	1118			
Db	143	atctaggctcacatgcaaacccccccaccttcaccaggtttcaagaacaatatgctgcctaagctc	202			
Oy	1119	ACCATGGCTCTGCGAGCGCTTGACCTCTGGGAGGCTCAAGATGTCCTCCTGCTCAGTCTC	1178			
Db	203	cgaagaagcttgtagcttcacagcatgatgccacacacactgtaat-ttttgttatltt-t	260			
Oy	1179	CTA-GTAGCTTGGAATCAACAMGAGGGGCCACACACCTGACTTAATCTTTTGTGTTTGT	1237			
Db	261	ttagtagaaggagggtttcacacatgtttggccaagctgtctagaaaccccaggatca	317			
Oy	1238	TGTGTAAGAATGAGCAATTCGCGCAATGTTGTATACAGGCTGCTCAAACTCTTAGGTCA	1294			
LOCUS	5	C06339	514 bp	mRNA	EST	11-OCT-1996
DEFINITION	similar to none.					
ACCESSION	C06339					
NID	g1503115					

KEYWORDS EST(expressed sequence tag).
SOURCE Homo sapiens human pancreatic islets cDNA to mRNA, clone_11b:Human pancreatic islet cDNA clone:hbc5534.
ORGANISM Homo sapiens
 Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Carnivora; Homidae; Homo.

REFERENCE 1 (sites)
AUTHORS Takeda, J.
TITLE not determined
JOURNAL Unpublished (1996)
REFERENCE 2 (bases 1 to 514)
AUTHORS Takeda, J.
TITLE Takeda, J.
JOURNAL Direct Submission
 Submitted (08-MAY-1996) to the DDBJ/EMBL/GenBank databases. Jun Takeda, Institute for Molecular and Cellular Regulation, Gunma University, Dept. of Molecular Medicine, 3-39-15 Shova-machi, Maebashi 371, Japan (E-mail:j.takeda@eweb.sj.gunma-u.ac.jp, Tel:81-272-20-8856, Fax:81-272-20-8896)
 Project: 'Human pancreatic islet cDNA'
 vector-Lambda ZAPII
 primer-T7 primer
 RsiteI=Eco RI
 Rsite2=Xho I

COMMENT mRNA was prepared from normal adult human islets. cDNA was directionally synthesized from the Xho I in the vector to the EcoRI site.
 cDNA was size fractionated to remove sequences <1000 bp in size.

FEATURES
 source
 1..514
 /organism="Homo sapiens"
 /clone="hbc5534"
 /clone_11b="Human pancreatic islet cDNA"
 /tissue_type="human pancreatic islets"
 113 a 158 c 96 g 143 t 4 others

BASE COUNT
 113 a 158 c 96 g 143 t 4 others

ORIGIN
 Query Match 7.5%; Score 106; DB 91; Length 514;
 Best Local Similarity 76.1%; Pred. No. 8.32e-105;
 Matches 175; Conservative 0; Mismatches 52; Indels 3; Gaps 3;

1 tttttttttttttagagcagagctgtcgtcgtgtcccaagcgtgagtcagtcagcagc 60
 1059 TTTT TTTT TTTT TTTT TTTT TGACAGGGGTCTACTCTGTACACCAAGCTGGAGTGCAGTGGCAC 1118
 61 atcttgcctactgcacagctccaccctccag-gttcaagcgcattctcctgcctcagctc 119
 1119 ACCAAGGCTCTCTGACGCTTGACCTCTGGGAGCTCAAGTGAATCCTCGCCTCAGTCTC 1178
 120 ccaagtagctagcagctacagatgacccgcacacacccccgcggttaattttttttgattt- 178
 1179 CTA-GTACTCTGGAATCAACAGGAGGACACACACCTGACCTAACTTTTGTGTTTGT 1237
 179 tttagatagatggaggtttaccatgtttgacagagtgatgtcagctctt 228
 1238 TTGCTAAGATGGCATTTTCGCCAATGTTGTACAGGCTGCTCAAACTCCT 1287

LOCUS 6 AA226159 465 bp mRNA EST 01-MAY-1997
DEFINITION ncl0h07.1 NCI CGAP P11 Homo sapiens cDNA clone 1910 similar to contains Alu repetitive element.
ACCESSION AA226159
KEYWORDS AA226159
KEYWORDS g1847475
KEYWORDS EST.
ORGANISM human.
 Homo sapiens
 Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Carnivora; Homidae; Homo.

REFERENCE 1 (bases 1 to 465)
AUTHORS NCI-CGAP.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL
Tumor Gene Index
Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui,
M.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html
For additional information: Robert Strausberg, Ph.D.
Phone: (301) 496-1550
E-mail: Robert.Strausberg@nih.gov
Seq primer: -28ml3 rev1.57 from Amersham
High quality sequence stop: 320.

FEATURES

Source

Location/Qualifiers

1..465

/organism="Homo sapiens"
/note="Vector: PAMP10; Site.1: Not; Site.2: EcoRI; 1st
strand cDNA was primed with oligo(dt)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected, histologically normal
prostate epithelial cells. Double-stranded cDNA was
ligated to EcoRI adaptors, 5 cycles of PCR applied to the
cDNA with an adaptor-specific primer, and the resulting
PCR product subcloned into PAMP10 by the UDG-cloning
method (Life Technologies). Average insert size is 600
bp. NOTE: Not directionally cloned. This library was
constructed by David Kitzman.
/clone="1910"
/clone_1lb="NCI CGAP Pr1"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
<1..>465

BASE COUNT

130 a 86 c 133 g 116 t

ORIGIN

mRNA

Query Match 7.4%; Score 105; DB 71; Length 465;
Best Local Similarity 74.0%; Pred. No. 2,13e-103;
Matches 185; Conservative 0; Mismatches 62; Indels 3; Gaps 3.

Db 217 gggggggaatcattgagccatgagttcaagacatgctgsggcaacatagtgagaccctgt 276
|||||
Cp 1306 GGGGAGGCCAAAGTAACCTAGAGATTGAGACCCAGCTGTACACAACTGGGAAATGCCAT 1247
Db 277 ctctacaaaataacaaaataaaaataatgactaggtgtgtgtgcatgcacctgtagtcg 336
|||||
Cp 1246 CTTTACCAAACA-AAAAAATAAAGATTAGACAGGTGTGGGGCCCTCTTGTAAGTTC 1188
Db 337 agctgccttgggaagcctgaagcgggaaggaattgctcag-tccaggaattcaagctcag 395
|||||
Cp 1187 AGCTACTAGG-AGACTAGAGCAGAGCATCACTTGACCTCCCAAGAGTCAAGCTGCAGA 1129
Db 396 aagctatgatcatcaccatcaccaccacattgggtgacagaagaagccctgtctccaaa 455
|||||
Cp 1128 GAGCCATGAGTGTCCACTGCCTCCAGCCTGGGTGACAGAGTGAGACCTGTCAAAAAA 1069
Db 456 aaaaaaaaaa 465
|||||
Cp 1068 AAAAAAAAAA 1059

RESULT 7

LOCUS AA229986 361 bp mRNA EST 01-MAY-1997

DEFINITION nc37405.g1 NCI CGAP Pr2 Homo sapiens cDNA clone 4450 similar to

ACCESSION AA229986

g1852280 71D

KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1 (bases 1 to 361)
AUTHORS	NCI-CCAP.
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CCAP).
JOURNAL	Tumor Gene Index
COMMENT	Unpublished (1997)

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui,
M.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/FLNT at:
www.bio.lnl.gov/dbir/image/image.html
For additional information: Robert Strausberg, Ph.D.
Phone: (301) 496-1550
E-mail: Robert.Strausberg@nih.gov
Seq primer: -41ml3 fwd, ER from Amersham
High quality sequence stop: 306.

FEATURES
SOURCE

Location/Qualifiers
1..361

MRNA	COUNT	ORIGIN
1	1	1
2	1	2
3	1	3
4	1	4
5	1	5
6	1	6
7	1	7
8	1	8
9	1	9
10	1	10
11	1	11
12	1	12
13	1	13
14	1	14
15	1	15
16	1	16
17	1	17
18	1	18
19	1	19
20	1	20
21	1	21
22	1	22
23	1	23
24	1	24
25	1	25
26	1	26
27	1	27
28	1	28
29	1	29
30	1	30
31	1	31
32	1	32
33	1	33
34	1	34
35	1	35
36	1	36
37	1	37
38	1	38
39	1	39
40	1	40
41	1	41
42	1	42
43	1	43
44	1	44
45	1	45
46	1	46
47	1	47
48	1	48
49	1	49
50	1	50
51	1	51
52	1	52
53	1	53
54	1	54
55	1	55
56	1	56
57	1	57
58	1	58
59	1	59
60	1	60
61	1	61
62	1	62
63	1	63
64	1	64
65	1	65
66	1	66
67	1	67
68	1	68
69	1	69
70	1	70
71	1	71
72	1	72
73	1	73
74	1	74
75	1	75
76	1	76
77	1	77
78	1	78
79	1	79
80	1	80
81	1	81
82	1	82
83	1	83
84	1	84
85	1	85
86	1	86
87	1	87
88	1	88
89	1	89
90	1	90
91	1	91
92	1	92
93	1	93
94	1	94
95	1	95
96	1	96
97	1	97
98	1	98
99	1	99
100	1	100

<1.,>361		
80 a	89 c	78 g 114 t

	Query Match	7.38	Score 104;	DB 722	Length 361;
	Best Local Similarity 75.88;		Pred. No. 5,43e-102;		
	Matches 175;	Conservative 0;	Mismatches 53;	Indels 3;	Gaps 3
Db	107	acattctcttttttttgacagcgagtttcacgc-gtgcaccagctgtagtgcatagctg	165		
Oy	1057	ACTTTTCTTTTCTTTTGTGACAGGGCTCTACCTCTCACACCAGGCTGAGTGCATGTGCA	1116		
Db	166	caatctggctgcacccgagcctccaccctccgg-gtttaattgattctctgccttaagcc	224		
Oy	1117	CCACCATGGCTCTGCGACGCTTGACCTCTTGAGACTCAAGTGAATCTCTCGCTCAATC	1176		
Db	225	tttcaagaagctgtagtcttcagcagcatgcccgcacatgctcgtacataaatttttttgat	284		
Oy	1177	T-CCCTAGAGCTGCACTACAAGAGGAGGCCACCCACACTGACTACTTGTGTTTTT	1235		
Db	285	ttttgtagagagcgagcttccatctgtgtgtaagcgctgagctcggaatactc	335		
Oy	1236	GTTTGTAAAGATGGCAATTTGCCCAATGTGTACAGCTGGTCTCAAACTCC	1286		

RESULT	8	AA326121	367 bp	EST	20-APR-1997
LOCUS		EST229525	Cerebellum II	Homo sapiens cDNA 5'	end similar to EST
DEFINITION			containing Alu repeat.		
ACCESSION		AA326121			
NID		91978609			
KEYWORDS		EST.			
SOURCE		human.			
ORGANISM		Homo sapiens			

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE AUTHORS

1 (bases 1 to 367)
Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,

TITLE	Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
JOURNAL	Nature 377 (6547 Suppl.), 3-174 (1995)
MEDLINE	96026280
COMMENT	

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
96026280

FEATURES

Location/Qualifiers
1..367

ORIGIN	BASE COUNT	mRNA
--------	------------	------

<1..>	367				
48 a	113	c	70	g	130
					t
					6 others

	Query March	7.38;	Score 103;	DB 36;	Length 367;
	Best Local Similarity	74.18;	Pred. No. 1.38e-100;		
	Matches 172; Conservative	0;	Mismatches 57;	Indels 3;	Gaps 3
Db	45	ccttcctccttctctccagatgtagtcgcctctccactcagcgmgagaaacagtggcgc	104		
Oy	1058	CTTTTCTTTTTTTTTTTTACAGGGGCTCACCCTGCACCCAGCGGTGGATGTGGCAC	1117		
Db	105	catctaaggctcactgcaacctccacctcngnt-tcaagnattctctcyctccanoc	163		
Oy	1118	CACCAATGGCTCTTCGACACCTTGACCTCTGGAGACTCAAGTAGATCCTCCCTCACTCT	1177		
Db	164	cocgagtgctgggataacaagtgtctggcaccaatgctgctgaatttgtatttttc	223		
Oy	1178	CCT-AGTAGCTGGAACTNCAAGAAGGCCACCAOACTGACTACTATTGTTTTGTTTTG	1236		

Contact: Wilson RK
WashU-Merck EST Project
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAE Consortium (info@mae.llnl.gov) for further information.
Insert Length: 751 Std Error: 0.00
Seq primer: mod.RegA+T
High quality sequence stop: 340

Location/Qualifiers
1..384

```

/organism="Homo sapiens"
/note="Organ: Liver and Spleen; Vector: pRT3D (Pharmacia)
With a modified polylinker. Site.1: Pac I; Site.2: Eco RI.
This is a subtracted version of the original Soares fetal
liver spleen INFUS library. 1st strand cDNA was primed
with a Pac I - 019g(ct) primer [5',
AATCGAGAGATTAATTAATGAAGCTTTTATTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pRT3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Felma Bonaldo."
/clone="148391"
/clone_id="Soares fetal liver spleen INFUS S1"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/complement("<1..384)
/db_xref="GDB:1326861"
95 c 82 g 116 t 1 others

```

7.3%; Score 104; DB 69; Length 384;
 Identity 72.5%; Pred. Mol. 5.43e-102;
 Descriptive 0; Mismatches 71; Indels 4; Gaps 4;

TACTCTGTCACCCAGGCTGGAGTGCAGTGGCACACAC-CATGGCTTCTGC 1133
 |||||
 TACTCTGTCACCCAGGCTGGAGTGCAGTGGCACACAC-CATGGCTTCTGC 62
 |||||
 TACTCTGTCACCCAGGCTGGAGTGCAGTGGCACACAC-CATGGCTTCTGC 121
 |||||
 TACTCTGTCACCCAGGCTGGAGTGCAGTGGCACACAC-CATGGCTTCTGC 1192
 |||||
 TACTCTGTCACCCAGGCTGGAGTGCAGTGGCACACAC-CATGGCTTCTGC 1192
 |||||

gagcgcttcagcancigtgcctggc 273
GACATGAATCGCCAGCCTCGGC 1345

ttgccacgctggtcttgaatccctggtctaagtgatccfccccaaagt 240
TTTTCNAGAAGCTGGTCTCAAACTCCTAAGGTCTACTTTGGCTCCCAAAGTG 1312

cccccaactgcalccgcgtacaaccttttggttttttt-agaatatggg 180
GCCCACACCACTGACNACCTTTTTGTTTTGTTGGTAAAGATGCCA 1252Z

426 bp mRNA EST 21-APR-1997
Brain IV Homo sapiens CDNA similar to EST containing Alu
ens

1 to 426).
D., Keriavavage, A.R., Fleischmann, R.D., Fuldner, R.A.,
Lee, N.H., Kirthness, E.F., Weinstock, K.G., Gocayne, J.D.,
Sutton, G., Blake, J.A., Brandon, R.C., Man, Wal, C.,
Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
Ged, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S.,
Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,
Kelley, J.C., Liu, L.-I., Marmorens, S.M., Merrick, J.M.,
Llanquanes, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R.,
Spriggs, T.A., Uterback, T.R., Weidman, J.F., Wiley,
D.P., Cao, L., Cepeda, M.C., Coleman, T.A., Collins, E.J.,
Feng, D.-F., Ferlie, A., Fischer, C., Haatling, G.A.,
Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,
Kunisch, C., Hungjun, J., Li, H., Meltsner, P.S., Olsen, H.,
L., Welty, F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M.,
P.M.J., Pennon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,
and Venter, J.C.

Assessment of human gene diversity and expression patterns
on 83 million nucleotides of cDNA sequence
(6347 Suppl), 3-174 (1995)

EST01568 THC110460

For further information, contact Dr. Kerstin Lawitzky, Institute for Genomic Research, Local Center Drive, Rockville, MD 20850 USA
E-mail: kerstin.lawitzky@tigr.org
Tel: 301-251-6990/699056
Fax: 301-251-6994/699423
Web: <http://www.tigr.org/rdb/ngi/ngi.html>
TIGR: M13-21.

```

Location/Qualifiers
1..426
/organism="Homo sapiens"
/note="Organ: Brain; Vector: plusescript SK-; Site:1"
ECORI; Site:2: Ecori"
/clone_lib="Brain IV"
/sex="female"
/dev_stage="fetus, 17-18 wks"
<1...>426
122 c 97 g 108 t 2 others
7.3%, Score 104, DB 62, Length 426;

```


mRNA <1...>428
 BASE COUNT 101 a 124 c 84 g 114 t 5 others
 ORIGIN
 Query Match 7.2%; Score 102; DB 71; Length 428;
 Best Local Similarity 73.6%; Pred. No. 3,49e-99;
 Matches 181; Conservative 0; Mismatches 62; Indels 3; Gaps 3;
 Db 1 ttttttttggagaagaggtctgtctgtcaccgaagtggatgcatgatca 60
 |||||||
 QY 1063 ttttttttttttggaaagggtctgactctgtccaccagagtgacacacaca 1122
 |||||||
 Db 61 tagctctcgtcagcttgaaactctctgga-cgtgaacgaatcccccagctccca 119
 |||||
 QY 1123 tggctctctgcagccttgatgacctctggagacctaaatgattctctgctcacttcccta- 1181
 |||||
 Db 120 gcaagctcgactacagtgatgcacacacacactagctaatttttattgtat 179
 |||||
 QY 1182 gtagctggaaactacaggaaggccacacacactgactgacatctttttgtttg- 1240
 |||||
 Db 180 gttagagacaggtctcaactatgttgcacagctagtcttgaaactaatggtccaatgat 239
 |||||||
 QY 1241 gtaaaagatggcattttgcgcatgtgtacagagctggctcaaaactcttaggttgg 1300
 |||||||
 Db 240 cctccc 245
 |||||||
 QY 1301 cctccc 1306
 |||||||
 RESULT 13
 LOCUS AA234890 468 bp mRNA EST 03-MAR-1997
 DEFINITION zr77h10.r1 Soares NHMNP St.Homo sapiens cDNA clone 669475 5'
 similar to contains Alu repetitive element.
 ACCESSION AA234890
 NID g1859401
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 468)
 Hollier,L., Clark,N., Dubnue,T., Elliston,K., Hawkins,M.,
 Parsons,M., Holtman,M., Kueba,T., Le,M., Lennon,G., Marra,M.,
 Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
 Tivavaas,K.E., Waterston,R., Williamson,A., Woldmann,P. and
 Wilson,R.
 The WashU-Merck EST Project
 Unpublished (1995)
 TITLE
 JOURNAL
 COMMENT
 Contact: Wilson RK
 WashU-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LMLT; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 432.
 Location/Qualifiers
 1..468
 /organism="Homo sapiens"
 /note="Organ: mixed (see below); Vector: pUT73D-Pac
 (Pharmacia) with a modified polylinker; Site_1: Not 1;
 Site_2: Eco RI; Equal amounts of plasmid DNA from three
 normalized libraries (melanocyte 2NbHV, pregnant uterus
 NBHV1, and fetal heart NBHV19W) were mixed, and ss circles
 were made in vitro. Following HAP purification, this DNA
 was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from pools of
 5,000 clones from the same 3 libraries. The pools
 consisted of I.M.A.G.E. clones 260232-265223,

[illegible]

	RESULT	14	C15363	351 bp	mRNA	EST	03-SEP-1996
	LOCUS						
	DEFINITION		Human fetal brain cDNA 5'-end	GEN=137E09.			
	ACCESSION		C15363				
	NID		g1570070				
	KEYWORDS		EST(expressed sequence tag); Human fetal brain.				
	SOURCE		Homo sapiens fetus brain cDNA to mRNA, clone:137E09.				
	ORGANISM		Homo sapiens				
			Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;				
			Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae;				
	REFERENCE		Homo				
	AUTHORS		1 (sites)				
			Fujiwara,T., Hirano,H., Hishigaki,H., Horie,M., Kawai,A., Kuga,Y.,				
			Kyushiki,I., Nagata,M., Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y.,				
			Shinomiyama,H., Suzuki,M., Takachi,A., Takeda,S., Watanabe,T.,				
			Maekawa,H., Nakamura,Y. and Takahashi,E.				
	TITLE		Otsuka cDNA project				
	JOURNAL		Unpublished (1996)				
	REFERENCE		2 (bases 1 to 351)				
	AUTHORS		Fujiwara,T.				
	TITLE		Direct Submission				
	JOURNAL		Submitted (13-MAY-1996) to the DDBJ/EMBL/Genbank databases. Tsutomu				
			Fujiwara, Otsuka GKN Research Institute,Otsuka Pharmaceutical				
			Co.,Ltd: 463-10 Kagasuno Kawmuchi-cho, Tokushima, Tokushima 771-01,				
			Japan (Tel:+81-886-65-2888, Fax:+81-886-37-1035)				
	FEATURES		Location/Qualifiers				
	SOURCE		1..351				
			/organism="Homo sapiens"				
			/clone="137E09"				
			/dev_stage="fetus"				
			/tissue_type="brain"				
BASE COUNT		90 a	86 c	67 g	98 t	10 others	
ORIGIN							
Query Match		7.1%;	Score 100;	DB 93;	Length 351;		
Best Local Similarity		76.2%;	Pred. No. 2,21e-96;				
Matches 176;	Conservative	7;	Mismatches 41;	Indels 7;	Gaps 4;		
Db	116 bbbtttyttttttttaagacagagtctcactctgttcacccaggctggatgcagtggcgc	175					
	::::: :						
Yr	1058 ctttttttttttttttacacgggcttcaccctctgctcacccacggctggatgcacgtggcgc	1117					

[illegible]

LOCUS	15	RESULT
AA195351	379 bp	mRNA
zr3g312.s1	Soares NHMPu S1	Homo sapiens CDNA clone 665254.3
DEFINITION	similar to contains Alu repetitive element;.	
Accession	AA195351	

EST. human.
Homo sapiens

REFERENCE AUTHORS

TITLE	The WashU-Merck EST Project
JOURNAL	Unpublished (1995)
COMMENT	

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewartson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAE Consortium (info@image.llnl.gov) for further information.
High quality sequence stop: 271.

FEATURES
SOURCE

1..379

/organism="Homo sapiens"
/note="Organ: mixed (see below); Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBH4, pregnant uterus
NBH9U, and fetal heart NBH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
/clone="665254"

	complement(<1..>379)
mRNA	75 a 103 c 87 g 114 t
BASE COUNT	
ORIGIN	

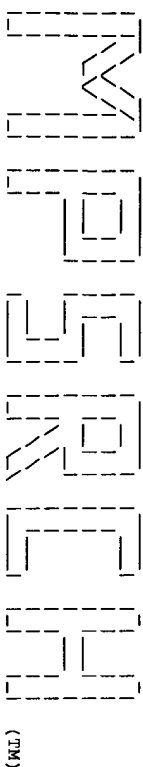
Query Match	7.1%;	Score 100;	DB 2;	Length 379;
Best Local Similarity	78.5%;	Pred. No. 2.21e-96;		
Matches	186;	Conservative	0;	Mismatches 44;
			Indels	7;
			Gaps	4;

```
Db      3 ttttttttttctgtgagacagtctgtcctgtgcaccaggctgtgaatggcgctg 62
        |||||
QY    1059 ttttttttttttgacagggtccactgtcacccaggctggatgacatggcacc 118
```

[illegible]

Search completed: Wed Mar 25 13:54:24 1998
Job time : 470 secs.

THIS PAGE BLANK (USPTO)



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligentics, Inc.

Mpsrch_un n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Wed Mar 25 12:55:36 1998; Maspar time 1516.90 Seconds
1325.012 Million cell updates/sec
Mar output not generated.

Title: >US-08-236-918A-7
Description: (1-1415) from 5674704.seq
Perfect Score: 1415
N.A. Sequence: 1 AGTGGAAAGTCTCCGGCAG.....TGANAAAAAAAAAAAAA 1415
Comp: TCACCTTCAAGAGCCGTC.....ACATATTTTATTTTATTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0
Searched: 430261 seqs, 710217276 bases x 2

Post-processing: Minimum Match 08
Listing first 45 summaries

Database: emb1-new7
1.BCT 2.FUN 3.GEN 4.HTG1 5.HTG2 6.HTG3 7.HTG4 8.HUM1
9.HUM2 10.HUM3 11.INV1 12.INV2 13.ORG 14.MAM 15.VRT
16.PLN 17.PRO1 18.PRO2 19.ROD 20.SYN 21.UNC 22.VIR
Database: genbank101
23.BCT1 24.BCT2 25.BCT3 26.BCT4 27.BCT5 28.BCT6 29.BCT7
30.BCT8 31.BCT9 32.BCT10 33.BCT11 34.BCT12 35.BCT13
36.GEN1 37.GEN2 38.GEN3 39.GEN4 40.GEN5 41.GEN6 42.HTG1
43.HTG2 44.HTG3 45.HTG4 46.HTG5 47.INV1 48.INV2 49.INV3
50.INV4 51.INV5 52.INV6 53.INV7 54.INV8 55.INV9 56.INV10
57.INV11 58.INV12 59.MAM1 60.MAM2 61.MAM3 62.VRT1
63.VRT2 64.VRT3 65.VRT4 66.PAT1 67.PAT2 68.PAT3 69.PAT4
70.PAT5 71.PAT6 72.PAT7 73.PHG 74.PLN1 75.PLN2 76.PLN3
77.PLN4 78.PLN5 79.PLN6 80.PLN7 81.PLN8 82.PLN9 83.PLN10
84.PLN11 85.PLN12 86.PRI1 87.PRI2 88.PRI3 89.PRI4
90.PRI5 91.PRI6 92.PRI7 93.PRI8 94.PRI9 95.PRI10
96.PRI11 97.PRI12 98.PRI13 99.PRI14 100.PRI15 101.PRI16
102.PRI17 103.ROD1 104.ROD2 105.ROD3 106.ROD4 107.ROD5
108.ROD6 109.ROD7 110.ROD8 111.ROD9 112.STR 113.SYN
114.UNA

Database: genbank101
115.VRL1 116.VRL2 117.VRL3 118.VRL4 119.VRL5 120.VRL6
121.VRL7 122.VRL8 123.VRL9 124.VRL10 125.VRL11
Database: genbank-new7
126.BCT 127.GEN 128.HTG1 129.HTG2 130.INV 131.MAM
132.VRT 133.PHG 134.PLN1 135.PLN2 136.PRI1 137.PRI2
138.ROD 139.SYN 140.UNA 141.VRL
Database: u-emb151_101
142.part1 143.part2

Statistics: Mean 11.786; Variance 6.409; scale 1.839

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1415	100.0	1415	94	HSU03397	Human receptor protel	0.00e+00
2	1378	97.4	1419	99	HUMILAX	Human activation depe	0.00e+00
3	233	16.5	2350	108	MSTC41BB	Mouse T-cell receptor	1.00e-145
4	131	9.3	940	101	HUMTRPM2A2	Human TRPM-2 protein	2.06e-69
5	129	9.1	158305	45	HS390013	Human DNA sequence **	5.94e-68
6	126	8.9	100234	128	HS341D10	Human DNA sequence **	9.12e-66
7	126	8.9	100234	5	HS341D10	Human DNA sequence **	9.12e-66
8	124	8.8	125492	44	HS167P19	Human DNA sequence **	9.12e-66
9	124	8.8	8339	8	HSAB1901	Human gene for L-hist	4.87e-65
10	125	8.8	32351	99	HSMULHDC	Human gene for L-hist	4.87e-65
11	125	8.8	100000	45	HSAC000012	*** SEQUENCING IN PRO	4.87e-65
12	125	8.8	100000	5	HSAC12	*** SEQUENCING IN PRO	4.87e-65
13	124	8.8	145194	5	HS339A18	Human DNA sequence **	2.60e-64
14	124	8.8	145194	128	HS339A18	Human DNA sequence **	2.60e-64
15	121	8.6	96625	89	HSAC000118	Human DNA sequence **	3.90e-62
16	121	8.6	112601	45	HS60N8	Human DNA sequence **	3.90e-62
17	121	8.6	137246	88	HS49J10	Human DNA sequence **	3.90e-62
18	122	8.6	166271	45	HS445C9	Human DNA sequence **	7.35e-63
19	122	8.6	224681	128	HS44N10	Human DNA sequence **	7.35e-63
20	122	8.6	225094	42	AC002044	*** SEQUENCING IN PRO	7.35e-63
21	122	8.6	225094	42	AC002044	*** SEQUENCING IN PRO	7.35e-63
22	122	8.5	225094	5	HSAC2044	*** SEQUENCING IN PRO	7.35e-63
23	120	8.5	4659	86	AC002212	Human DNA sequence **	2.07e-61
24	120	8.5	4659	6	HSAC2212	Human DNA sequence **	2.07e-61
25	120	8.5	27584	137	HSB5E3A	Human DNA sequence **	2.07e-61
26	120	8.5	27584	8	HSB5E3A	Human DNA sequence **	2.07e-61
27	120	8.5	27584	90	HSB5E3A	Human DNA sequence **	2.07e-61
28	120	8.5	100000	0	HSAC11	*** SEQUENCING IN PRO	2.07e-61
29	120	8.5	100000	45	HSAC000011	*** SEQUENCING IN PRO	2.07e-61
30	119	8.4	37027	88	HS0506ZB	Human DNA sequence **	1.10e-60
31	119	8.4	43952	94	HS0105G4	Human DNA sequence **	1.10e-60
32	119	8.4	76094	87	HS179D3A	Human DNA sequence **	1.10e-60
33	119	8.4	95425	91	HSJ506G21	Human DNA sequence **	1.10e-60
34	119	8.4	127079	87	HS214K23	Human DNA sequence **	1.10e-60
35	119	8.4	130679	44	HS21D12	Human DNA sequence **	1.10e-60
36	118	8.3	2802	89	HSAC000966	Human DNA sequence **	5.80e-60
37	118	8.3	2854	100	HUMN1F19AG	Human NFG genomic fra	5.80e-60
38	118	8.3	2897	92	HS181793	Human sapiens (sublon	5.80e-60
39	118	8.3	326	92	HS181575	Human sapiens (sublon	5.80e-60
40	118	8.3	13204	98	HDMGPIIB2	Human platelet glycop	5.80e-60
41	118	8.3	132098	88	HS46H23	Human DNA sequence **	5.80e-60
42	118	8.3	132645	129	AC002345	*** SEQUENCING IN PRO	5.80e-60
43	118	8.3	132645	6	HSAC2345	*** SEQUENCING IN PRO	5.80e-60
44	118	8.3	142395	89	HSAC000120	Human BAC clone RG161	5.80e-60
45	118	8.3	210672	44	HS179115	Human DNA sequence **	5.80e-60

ALIGNMENTS

RESULT	LOCUS	1	HSU03397	1415 bp	mRNA	PRI	15-NOV-1994
DEFINITION	Human Receptor protein 4-1BB mRNA, complete cds.						
ACCESSION	U03397						
NID	9571320						
KEYWORDS	human.						
SOURCE	human.						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 1415)						
AUTHORS	Algeron M.R., Smith,C.A., Tough,T.W., Davis-Smith,T., Armitage,R.J., Falk,B., Roux,E., Baker,E., Sutherland,G.R., Din,W.S. and Goodwin,R.G.						
TITLE	Molecular and biological characterization of human 4-1BB and its ligand						
JOURNAL	Eur. J. Immunol. 24 (9), 2219-2227 (1994)						

MEDLINE 94374434
 REFERENCE 2 (bases 1 to 1415)
 AUTHORS Alderson, M.
 TITLE Direct Submission
 JOURNAL Submitted (10-NOV-1993) Mark Alderson, Immunex Research and Development Corp., 51 University St., Seattle, WA 98101, USA

FEATURES
 source
 1..1415
 /organism="Homo sapiens"
 /tissue-type="peripheral blood"
 /cell-type="T cell"
 120..887
 /note="homology to the receptors for TNF and NGF; human homolog of murine T-cell receptor 4-1BB protein; transmembrane protein"
 /codon_start=1
 /product="4-1BB"
 /db_xref="PID:9571321"
 /translation="MNSCYNIIVATLILVNFERTSLDPCSNCPAGTFCDDNNMOI CSPSPNFSASGQRTCDICROCKGVTRKREKSTSNACDTPGPHGAGCSMC EODCKOGELRKKGKDCFGFNDRKRGKIPMTNCSLDKSYLVNGTKRDVYCGP SPADSPASSVTTPAPARBERGHSQIISFLALSTPLLLFLFFILIRFSVYKGRK KLITFKQFPMKPVQTTQEDGCSQFPEEEGGCEL"

misc_feature
 531..533
 /note="encodes potential glycosylation site"
 misc_feature
 564..566
 /note="encodes potential glycosylation site"
 misc_feature
 678..758
 /note="encodes transmembrane domain; amino acids 187-208"

BASE COUNT 385 a 332 c 333 g 365 t
 ORIGIN

Query Match 100.0%; Score 1415; DB 94; Length 1415;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 1415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db      1 agtggaaagttctccgagcccttgagatcgaagatgacatttgagaccagtaat 60
Qy      1 AGTGAAGTTCTCCGGGAGCCCTGAGATCTCAAGAGTACATTGTGTGAGACAGTAT 60

Db      61 ttgattaaatctctcttggaatcagcttgtagtatacatcactgtgccagattcata 120
Qy      61 TTGATTAAATCTCTTGGAATCAGCTTGTGATATCATCTGTGCCAGATTTCATCA 120

Db      121 tgggaaacagctgtacaaatagtagccactgtgtgtgtgtgtgtgtgtgtgtgtgtgt 180
Qy      121 TGGGAAACGCTGTACAAATAGTAGCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180

Db      181 caagatcatgcaagatcctgttagtaactgcccagctgtgatactctgtgataataca 240
Qy      181 CAAGATCATGCAAGATCCTGTAGTAATGCCCAGCTGTGATCTGTGATTAATACA 240

Db      241 ggaatcagatttgacgtccctgtctccaatagttcttcacagcgcaaggtggacaagga 300
Qy      241 GGAATCAGATTTGACGTCCTGTCTCCAAATAGTTTCCACGCGAGGTGACAAGGA 300

Db      301 cctggaacataagcagcgagctgttaaagtggtttcaagaccaggaagagtggttctcca 360
Qy      301 CCTGGAACATAGCAGCGAGCTGTAAAGTGTTTCAAGACCAAGAGAGTGTTCTCTCA 360

Db      361 ccagaatgcaagagtgtagtcaactccagaggtttcaactgctggggaggaagatcagca 420
Qy      361 CCAGAAATGCAAGAGTGTAGTCACTCCAGGCTTTCACCTGCTGGGGAGGATGACAGA 420

Db      421 tgtgtgaacagagattgttaacaaggltcaagaactgcacaaaaaagggtgtgaaagactgt 480
Qy      421 TGTGTGAACAGGATTGTAAACAAGGTCAGAACTGACAAAAAAGGTGTAAAGACTGT 480

Db      481 gcttggagacatttaacgatacagaacgtggaatctgttcgacccctggacaacactgtctt 540
Qy      481 GCTTGGAGACATTTAACGATACGAACGCGGATCTGTGACCCCTGGACAAACTGTTCT 540

Db      541 tggatggaagagctgtgctgtgtaagtggaggaagaggaagagctgtgctgtggacat 600

```

```

Qy      541 TGGATGAAAGTCTGTGCTTGTGATGGAGCAAGAGGAGACTGTGCTGTGACCAT 600

Db      601 ctcacgccgacctctctccggyagacatcctgtbaccgccctcctccgcgagaagc 660
Qy      601 CTCACGCCGACCTCTCTCCGGAGCATCTCTGTACCCCGCTCCCTCGAGAGAGC 660

Db      661 caggaacactctccgagatacatctctctctctctctctctctctctctctctctct 720
Qy      661 CAGGACACTCTCCGAGATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720

Db      721 tctgtctgttctctctcagctcgtctctctctctgtgtttaaaggggacagaacatcc 780
Qy      721 TCTGCTGTGTCTCTCTCAAGCTCGTCTCTGTGTAAACGGGGAGAAAGAAATCC 780

Db      781 tgtatattatcaaacacattatagagccagttcaactactaagaagaagatgact 840
Qy      781 TGTATATTATCAAAACACATTATAGAACACAGTCAAAACTACTCAAGAGAGATGACT 840

Db      841 gtagctgcagatttcagaagaagaagaagagatgtgaaactgtgaaatgaagtaac 900
Qy      841 GTAGCTGCCGATTTCCAGAGAGAGAGAGAGATGTGAATGTGAATGGAAGTCAAT 900

Db      901 aggcctgtggaacttctctgaaagaagaagaataatagtcacatccgtatcacagc 960
Qy      901 AGGCTGTGGGACTTCTTGAAGAGAGCAAGAAATATAGTATCGCTATCACAGC 960

Db      961 ttcaaaagcaagaacacacatcctacataaccagaatcccccacacagttctt 1020
Qy      961 TTTCAAAAGCAAGAACACATCTCATATATCCAGAGATTCGCCCAACACAGTCTTT 1020

Db      1021 tctaattccaatgagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1080
Qy      1021 TCTAATCCCAATGAGTGTGGCTTTTAAATGACACACTTTTGTGTGTGTGTGTGTGTGT 1080

Db      1081 gtctacactgttcaccagcgctgtagtgcaagtggaccacacatgctctcgcagctgt 1140
Qy      1081 GTCTACTGTGTACACCGAGGCGAGTGCAATGGACACCATGCTCTCGACACCTTG 1140

Db      1141 acctctggagagctcaagtgtatcctctcgtcctcagtcctctgtagctggaactacaaga 1200
Qy      1141 ACCTCTGGAGAGCTCAAGTGAATCCTCTCTCCTCAGTCTCTGTAGCTGGAACAGGA 1200

Db      1201 agggccacacaacactgaactatttggtttggttgtgtaagaatgagatttcgcca 1260
Qy      1201 AGGGCCACACACCTGAATTAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260

Db      1261 tgtgtacagcgtgtgtccaaactcctagttcacttggcctcccaagtgctggaatt 1320
Qy      1261 TGTGTACAGCGTGTGTCAAACTCTAGTTCACTTTGGCTCCCAAGTGTCTGGATT 1320

Db      1321 acagacatgaactgcccagcccgccaaataatgacacactttaaagaaacagacaga 1380
Qy      1321 ACAGACATGAAGTGCAGGCCCGCCAAATATATCACCACACTTTTAACAGAACAGACA 1380

Db      1381 tgaggacagagctgtgtgtataaaaaaagagagagagagagagagagagagagagag 1415
Qy      1381 TGAGGACAGAGCTGTGTATATAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1415

```

RESULT 2
 LOCUS HUM11AX 1419 bp mRNA PRI 28-FEB-1995
 DEFINITION Human activation dependent T cell mRNA, complete cds.
 ACCESSION L12964
 NID 9292237
 KEYWORDS cell surface receptor; nerve growth factor receptor; human 4-1BB
 SOURCE Homo sapiens cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Euteria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1419)
 AUTHORS Schwarz, H., Tuckwell, J. and Lotz, M.

RESULT	5	DNA	HTG	22-JAN-1997
LOCUS	HS390013	158305 bp		
DEFINITION	Human DNA sequence *** SEQUENCING IN PROGRESS ***			from clone 390013; HTGS phase 1.
ACCESSION	284469			
NID	G1806012			
KEYWORDS	HTGS; HTGS_PHASE1.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homindae; Homo. 1 (bases 1 to 158305)			
AUTHORS	Vaidin,M.			
TITLE	Direct Submission			
JOURNAL	Submitted (06-NOV-1996) Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: hammyr@sanger.ac.uk			
COMMENT	IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated with foreign sequence from E.coli, yeast, vector, phage etc. Order of segments is not known; 800 n's separate segments. Unfinished sequence: dj390013 Contig_ID: 00607 length: 5173 bp unfinished sequence: dj390013 Contig_ID: 01014 length: 1883 bp unfinished sequence: dj390013 Contig_ID: 00221 length: 1122 bp unfinished sequence: dj390013 Contig_ID: 00328 length: 1779 bp unfinished sequence: dj390013 Contig_ID: 00656 length: 6569 bp unfinished sequence: dj390013 Contig_ID: 00423 length: 2789 bp unfinished sequence: dj390013 Contig_ID: 00388 length: 1088 bp unfinished sequence: dj390013 Contig_ID: 00236 length: 2833 bp unfinished sequence: dj390013 Contig_ID: 00755 length: 3330 bp unfinished sequence: dj390013 Contig_ID: 00491 length: 2337 bp unfinished sequence: dj390013 Contig_ID: 00685 length: 1168 bp unfinished sequence: dj390013 Contig_ID: 00023 length: 1554 bp unfinished sequence: dj390013 Contig_ID: 00630 length: 2127 bp unfinished sequence: dj390013 Contig_ID: 00158 length: 3524 bp unfinished sequence: dj390013 Contig_ID: 00800 length: 1506 bp unfinished sequence: dj390013 Contig_ID: 00618 length: 1016 bp unfinished sequence: dj390013 Contig_ID: 00984 length: 2910 bp unfinished sequence: dj390013 Contig_ID: 00340 length: 6914 bp unfinished sequence: dj390013 Contig_ID: 00948 length: 1080 bp unfinished sequence: dj390013 Contig_ID: 01083 length: 2145 bp unfinished sequence: dj390013 Contig_ID: 00489 length: 3945 bp unfinished sequence: dj390013 Contig_ID: 00583 length: 1508 bp unfinished sequence: dj390013 Contig_ID: 01063 length: 2337 bp unfinished sequence: dj390013 Contig_ID: 00716 length: 6262 bp unfinished sequence: dj390013 Contig_ID: 00671 length: 812 bp unfinished sequence: dj390013 Contig_ID: 00231 length: 859 bp unfinished sequence: dj390013 Contig_ID: 00191 length: 1457 bp unfinished sequence: dj390013 Contig_ID: 00683 length: 918 bp unfinished sequence: dj390013 Contig_ID: 00746 length: 1262 bp unfinished sequence: dj390013 Contig_ID: 00435 length: 2095 bp unfinished sequence: dj390013 Contig_ID: 00915 length: 1658 bp unfinished sequence: dj390013 Contig_ID: 00525 length: 2032 bp unfinished sequence: dj390013 Contig_ID: 00564 length: 2950 bp unfinished sequence: dj390013 Contig_ID: 00997 length: 3549 bp unfinished sequence: dj390013 Contig_ID: 00014 length: 3642 bp unfinished sequence: dj390013 Contig_ID: 00401 length: 2137 bp unfinished sequence: dj390013 Contig_ID: 00759 length: 750 bp unfinished sequence: dj390013 Contig_ID: 00876 length: 7670 bp unfinished sequence: dj390013 Contig_ID: 00725 length: 1849 bp unfinished sequence: dj390013 Contig_ID: 00658 length: 3695 bp unfinished sequence: dj390013 Contig_ID: 00998 length: 2096 bp unfinished sequence: dj390013 Contig_ID: 00528 length: 1255 bp unfinished sequence: dj390013 Contig_ID: 00108 length: 1028 bp unfinished sequence: dj390013 Contig_ID: 00702 length: 892 bp unfinished sequence: dj390013 Contig_ID: 00518 length: 2133 bp unfinished sequence: dj390013 Contig_ID: 00798 length: 2759 bp unfinished sequence: dj390013 Contig_ID: 00941 length: 2589 bp unfinished			

```

sequence: dj390013 Contig_ID: 01019 Length: 3698 bp.
***
*** WARNING: Phase 1 High Throughput Genome Sequence ***
*** This sequence is unfinished. When sequencing is complete,
*** the sequence data presented in this record will be replaced
*** by a single finished sequence with the same accession number.
Location/Qualifiers
1..158305
/organism="Homo sapiens"
/clone="390013"
/chromosome="X"
BASE COUNT 35945 a 24148 c 23618 g 36900 t 37694 others
ORIGIN
Query Match 9.1%; Score 129; DB 45; Length 158305;
Best Local Similarity 78.3%; Pred. No. 5,946-68;
Matches 188; Conservative 0; Mismatches 50; Indels 1; Gaps 1;
Db 133120 ttctcttttttgagacagggccttgctctgtcactcactgagctatattgtatttctttg 133179
Oy 1061 TTTTCTTTTGTGTTTGTGACAGGGGTCCTGCTGTCACCCAGGCTGGAGTGACAGGGCACCAC 1120
Db 133180 catgattcactgcagcctctgcacatcccaaggctcaagtgatctctccacatcagccccc 133239
Oy 1121 CATGCGCTCTGCGACACCTTGACCTCTGGAGACTCAAGAGATCTCTGCGCTCAGCTCTCT 1180
Db 133240 gagtcgagggagctcaagtcgtctgcacatcactgagctatattgtatttctttg 133299
Oy 1181 -AGTACGCTGGACACTACAAAGAAAGAGGCCACACACACTGACTACTTCTTTTGTGTTT 1239
Db 133300 ngtagagatggggtttccaccatattgscagagctgtgtcctaactcctcgtggtca 133354
Oy 1240 GGTAAAGATGGCATTTCCGCAATGTGTACAGGCTGCTCTCAACTCTAAGGTCA 1294

RESULT 6
LOCUS HS341D10 100234 bp DNA HTG 22-JUL-1997
DEFINITION Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone
341D10: HTGS phase 1.
ACCESSION 297985
NID 92274953
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae;
Homo.
1 (bases 1 to 100234)
REFERENCE 1
AUTHORS Buck,D.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-1997) Wellcome Trust Genome Campus, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humangid@anger.ac.uk Clone requests: clonerequest@anger.ac.uk
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence. Work on the sequence is in progress
and
the release of this data is based on the understanding that the
sequence may change as work continues. The sequence may be
contaminated
with foreign sequence from E.coli, yeast, vector, phase etc. Order
of segments is not known: 800 n/s separate segments. Unfinished
sequence: dj341D10 Contig_ID: 00214 Length: 49412 bp Unfinished
sequence: dj341D10 Contig_ID: 00359 Length: 50022 bp.
***
*** WARNING: Phase 1 High Throughput Genome Sequence ***
***
*** This sequence is unfinished. When sequencing is complete,
*** the sequence data presented in this record will be replaced
*** by a single finished sequence with the same accession number.
Location/Qualifiers
1..100234
/organism="Homo sapiens"

```

```

BASE COUNT      29376 a 20917 c 21419 g 27472 t 1050 others
ORIGIN
Query Match          8.9%; Score 126; DB 128; Length 100234;
Best Local Similarity 79.6%; Pred. No. 9,12e-66;
Matches 183; Conservative 0; Mismatches 45; Indels 2; Gaps 2

Db 74992 tttttttttttttagatgagatcctctgtcacccaggctgagtgacgatg 75051
Oy 1059 TTTT'TTTTTTTTTTTTGACAGGGCTCTCACTGTGTACCAGGCTGAGTGCATG 1118
Db 75052 atcctggtcactgcgaacctctgctcccgg- gtccaagtatttcctgctcaagctc 75110
Oy 1119 ACCAAGGCTCTCTCGAGGCTTCGACTCTCGGGAGCCTCAAGTGTCTCCCTCGCTCATGTC 1178
Db 75111 ctgagtagctcgagatcacagcatgtaccacaagccccggctaatttgttatcttagt 75170
Oy 1179 CT-ACTACTCTGAATATAACAAGAAAGGCCACACACTGTACTTAATT'TTTTGT' 1237
Oy 5171 agagctggagatggagattccacatgttgccagsgtgggtctcaaacgct 75220
Oy 1238 TTGTGAAGAAGATGGCATTTGGCATTTGTGTACAGGCTGTGTCTCAACTCTCT 1287

RESULT
ID HS341D10 standard; DNA; HTG: 100234 BP.
AC Z97985;
NT e1056721
DT 22-JUL-1997 (Rel. 52, Created)
DT 22-JUL-1997 (Rel. 52, Last updated, Version 1)
DE Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone 341D10
KW HTG; PHASE1.
OS Homo sapiens (human)
OC Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; CC Homo.
RN [1]
RP 1-100234
RA Buck D.;
RL Submitted (20-JUL-1997) to the EMBL/GenBank/DBJ databases.
RL Wellcome Trust Genome Campis, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
CC IMMORTALT. This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and
CC the release of this data is based on the understanding that the sequence may change as work continues. The sequencing may be contaminated
CC with foreign sequence from E.coli, yeast, vector, phage etc.
CC Order of segments is not known; 800 n's separate segments.
CC Unfinished sequence: d0341D10 Contig_ID: 00214 Length: 49412 bp
CC Unfinished sequence: d0341D10 Contig_ID: 00359 Length: 50022 bp
FH Key Location/Qualifiers
FT source 1..100234
FT FT /organism="Homo sapiens"
FT FT /clone="341D10"
FT FT /chromosome="X"
SQ Sequence 100234 BP; 29376 A; 20917 C; 21419 G; 27472 T; 1050 other;

Query Match          8.9%; Score 126; DB 5; Length 100234;
Best Local Similarity 79.6%; Pred. No. 9,12e-66;
Matches 183; Conservative 0; Mismatches 45; Indels 2; Gaps 2

Db 74992 tttttttttttttagagatgagatcctctgtcacccaggctgagtgacgatg 75051
Oy 1059 TTTT'TTTTTTTTTTTTGACAGGGCTCTCACTGTGTACCAGGCTGAGTGCATG 1118
Db 75052 atcctggtcactgcgaacctctgctcccgg- gtccaagtatttcctgctcaagctc 75110
Oy 1119 ACCAAGGCTCTCTCGAGGCTTCGACTCTCGGGAGCCTCAAGTGTCTCCCTCGCTCATGTC 1178
Db 75111 ctgagtagctcgagatcacagcatgtaccacaagccccggctaatttgttatcttagt 75170
Oy 1179 CT-ACTACTCTGAATATAACAAGAAAGGCCACACACTGTACTTAATT'TTTTGT' 1237
Oy 5171 agagctggagatggagattccacatgttgccagsgtgggtctcaaacgct 75220
Oy 1238 TTGTGAAGAAGATGGCATTTGGCATTTGTGTACAGGCTGTGTCTCAACTCTCT 1287

RESULT
ID HS341D10 standard; DNA; HTG: 100234 BP.
AC Z97985;
NT e1056721
DT 22-JUL-1997 (Rel. 52, Created)
DT 22-JUL-1997 (Rel. 52, Last updated, Version 1)
DE Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone 341D10
KW HTG; PHASE1.
OS Homo sapiens (human)
OC Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; CC Homo.
RN [1]
RP 1-100234
RA Buck D.;
RL Submitted (20-JUL-1997) to the EMBL/GenBank/DBJ databases.
RL Wellcome Trust Genome Campis, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
CC IMMORTALT. This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and
CC the release of this data is based on the understanding that the sequence may change as work continues. The sequencing may be contaminated
CC with foreign sequence from E.coli, yeast, vector, phage etc.
CC Order of segments is not known; 800 n's separate segments.
CC Unfinished sequence: d0341D10 Contig_ID: 00214 Length: 49412 bp
CC Unfinished sequence: d0341D10 Contig_ID: 00359 Length: 50022 bp
FH Key Location/Qualifiers
FT source 1..100234
FT FT /organism="Homo sapiens"
FT FT /clone="341D10"
FT FT /chromosome="X"
SQ Sequence 100234 BP; 29376 A; 20917 C; 21419 G; 27472 T; 1050 other;

Query Match          8.9%; Score 126; DB 5; Length 100234;
Best Local Similarity 79.6%; Pred. No. 9,12e-66;
Matches 183; Conservative 0; Mismatches 45; Indels 2; Gaps 2

Db 74992 tttttttttttttagagatgagatcctctgtcacccaggctgagtgacgatg 75051
Oy 1059 TTTT'TTTTTTTTTTTTGACAGGGCTCTCACTGTGTACCAGGCTGAGTGCATG 1118
Db 75052 atcctggtcactgcgaacctctgctcccgg- gtccaagtatttcctgctcaagctc 75110
Oy 1119 ACCAAGGCTCTCTCGAGGCTTCGACTCTCGGGAGCCTCAAGTGTCTCCCTCGCTCATGTC 1178
Db 75111 ctgagtagctcgagatcacagcatgtaccacaagccccggctaatttgttatcttagt 75170
Oy 1179 CT-ACTACTCTGAATATAACAAGAAAGGCCACACACTGTACTTAATT'TTTTGT' 1237
Oy 5171 agagctggagatggagattccacatgttgccagsgtgggtctcaaacgct 75220
Oy 1238 TTGTGAAGAAGATGGCATTTGGCATTTGTGTACAGGCTGTGTCTCAACTCTCT 1287

RESULT
ID HS341D10 standard; DNA; HTG: 100234 BP.
AC Z97985;
NT e1056721
DT 22-JUL-1997 (Rel. 52, Created)
DT 22-JUL-1997 (Rel. 52, Last updated, Version 1)
DE Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone 341D10
KW HTG; PHASE1.
OS Homo sapiens (human)
OC Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; CC Homo.
RN [1]
RP 1-100234
RA Buck D.;
RL Submitted (20-JUL-1997) to the EMBL/GenBank/DBJ databases.
RL Wellcome Trust Genome Campis, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
CC IMMORTALT. This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and
CC the release of this data is based on the understanding that the sequence may change as work continues. The sequencing may be contaminated
CC with foreign sequence from E.coli, yeast, vector, phage etc.
CC Order of segments is not known; 800 n's separate segments.
CC Unfinished sequence: d0341D10 Contig_ID: 00214 Length: 49412 bp
CC Unfinished sequence: d0341D10 Contig_ID: 00359 Length: 50022 bp
FH Key Location/Qualifiers
FT source 1..100234
FT FT /organism="Homo sapiens"
FT FT /clone="341D10"
FT FT /chromosome="X"
SQ Sequence 100234 BP; 29376 A; 20917 C; 21419 G; 27472 T; 1050 other;

Query Match          8.9%; Score 126; DB 5; Length 100234;
Best Local Similarity 79.6%; Pred. No. 9,12e-66;
Matches 183; Conservative 0; Mismatches 45; Indels 2; Gaps 2

Db 74992 tttttttttttttagagatgagatcctctgtcacccaggctgagtgacgatg 75051
Oy 1059 TTTT'TTTTTTTTTTTTGACAGGGCTCTCACTGTGTACCAGGCTGAGTGCATG 1118
Db 75052 atcctggtcactgcgaacctctgctcccgg- gtccaagtatttcctgctcaagctc 75110
Oy 1119 ACCAAGGCTCTCTCGAGGCTTCGACTCTCGGGAGCCTCAAGTGTCTCCCTCGCTCATGTC 1178
Db 75111 ctgagtagctcgagatcacagcatgtaccacaagccccggctaatttgttatcttagt 75170
Oy 1179 CT-ACTACTCTGAATATAACAAGAAAGGCCACACACTGTACTTAATT'TTTTGT' 1237
Oy 5171 agagctggagatggagattccacatgttgccagsgtgggtctcaaacgct 75220
Oy 1238 TTGTGAAGAAGATGGCATTTGGCATTTGTGTACAGGCTGTGTCTCAACTCTCT 1287

RESULT
ID HS341D10 standard; DNA; HTG: 100234 BP.
AC Z97985;
NT e1056721
DT 22-JUL-1997 (Rel. 52, Created)
DT 22-JUL-1997 (Rel. 52, Last updated, Version 1)
DE Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone 341D10
KW HTG; PHASE1.
OS Homo sapiens (human)
OC Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; CC Homo.
RN [1]
RP 1-100234
RA Buck D.;
RL Submitted (20-JUL-1997) to the EMBL/GenBank/DBJ databases.
RL Wellcome Trust Genome Campis, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
CC IMMORTALT. This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and
CC the release of this data is based on the understanding that the sequence may change as work continues. The sequencing may be contaminated
CC with foreign sequence from E.coli, yeast, vector, phage etc.
CC Order of segments is not known; 800 n's separate segments.
CC Unfinished sequence: d0341D10 Contig_ID: 00214 Length: 49412 bp
CC Unfinished sequence: d0341D10 Contig_ID: 00359 Length: 50022 bp
FH Key Location/Qualifiers
FT source 1..100234
FT FT /organism="Homo sapiens"
FT FT /clone="341D10"
FT FT /chromosome="X"
SQ Sequence 100234 BP; 29376 A; 20917 C; 21419 G; 27472 T; 1050 other;

Query Match          8.9%; Score 126; DB 5; Length 100234;
Best Local Similarity 79.6%; Pred. No. 9,12e-66;
Matches 183; Conservative 0; Mismatches 45; Indels 2; Gaps 2

Db 74992 tttttttttttttagagatgagatcctctgtcacccaggctgagtgacgatg 75051
Oy 1059 TTTT'TTTTTTTTTTTTGACAGGGCTCTCACTGTGTACCAGGCTGAGTGCATG 1118
Db 75052 atcctggtcactgcgaacctctgctcccgg- gtccaagtatttcctgctcaagctc 75110
Oy 1119 ACCAAGGCTCTCTCGAGGCTTCGACTCTCGGGAGCCTCAAGTGTCTCCCTCGCTCATGTC 1178
Db 75111 ctgagtagctcgagatcacagcatgtaccacaagccccggctaatttgttatcttagt 75170
Oy 1179 CT-ACTACTCTGA
```

Qy	1119	ACCAATGGCTCTCTGTGACGGCTTGACCTCTGCGGACGCTCAAGTGAATCTCTGCTGCCTCAAGTCTC	1178		
Db	75111	ctgagtagctcggagattacaggcatgtaccaccacgcccggcctaattttgtaatttagt	75170		
Qy	1179	CT-AGTACTGGAACTACAAAGAGGGCCACACACACTGACTAACTTTTGTTTTGT	1237		
Db	75171	agagctggagtgagattccaccattgtggccagcgtgtgtccaacgct	75220		
Qy	1238	TTGTAAAGATGGCATTTGCGCATGTTGTACAGGCTGGCTCTCAAACTGCT	1287		
LOCUS	8	HS167P19	125492 bp DNA	HTG	19-MAR-1997
DEFINITION		Human DNA sequence ***	SEQUENCING IN PROGRESS ***		from clone 167P19; HTGS phase 1.
ACCESSION					293014
NID					g1903191
KEYWORDS					HTG; HTGS_PHASE1.
SOURCE					human.
ORGANISM					Homo sapiens
REFERENCE					Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 125492)
AUTHORS					Vaudin,M.
TITLE					Direct Submission
JOURNAL					Submitted (06-NOV-1996) Wellcome Trust Genome Campus, Hinxton, Cambridge;shire.ac.uk Clone requests: clonequest@sanger.ac.uk
COMMENT					humbury@sanger.ac.uk Clone requests: clonequest@sanger.ac.uk IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated with foreign sequence from E.coli, yeast, vector, phage etc. Order of segments is not known; 800 n s separate segments. Unfinished sequence: d16167P19 Contig_ID: 00815 length: 974 bp Unfinished sequence: d16167P19 Contig_ID: 00002 length: 1262 bp Unfinished sequence: d16167P19 Contig_ID: 00391 length: 1106 bp Unfinished sequence: d16167P19 Contig_ID: 00346 length: 2190 bp Unfinished sequence: d16167P19 Contig_ID: 00822 length: 798 bp Unfinished sequence: d16167P19 Contig_ID: 00112 length: 1193 bp Unfinished sequence: d16167P19 Contig_ID: 00772 length: 1574 bp Unfinished sequence: d16167P19 Contig_ID: 00607 length: 2231 bp Unfinished sequence: d16167P19 Contig_ID: 00655 length: 898 bp Unfinished sequence: d16167P19 Contig_ID: 00241 length: 1020 bp Unfinished sequence: d16167P19 Contig_ID: 00464 length: 4127 bp Unfinished sequence: d16167P19 Contig_ID: 00734 length: 1144 bp Unfinished sequence: d16167P19 Contig_ID: 00012 length: 2889 bp Unfinished sequence: d16167P19 Contig_ID: 00075 length: 2137 bp Unfinished sequence: d16167P19 Contig_ID: 00652 length: 2620 bp Unfinished sequence: d16167P19 Contig_ID: 00262 length: 7559 bp Unfinished sequence: d16167P19 Contig_ID: 00742 length: 1380 bp Unfinished sequence: d16167P19 Contig_ID: 00252 length: 1684 bp Unfinished sequence: d16167P19 Contig_ID: 00899 length: 10303 bp Unfinished sequence: d16167P19 Contig_ID: 00044 length: 1720 bp Unfinished sequence: d16167P19 Contig_ID: 00993 length: 3193 bp Unfinished sequence: d16167P19 Contig_ID: 00887 length: 1021 bp Unfinished sequence: d16167P19 Contig_ID: 00653 length: 1730 bp Unfinished sequence: d16167P19 Contig_ID: 00107 length: 2209 bp Unfinished sequence: d16167P19 Contig_ID: 00240 length: 2802 bp Unfinished sequence: d16167P19 Contig_ID: 00054 length: 1333 bp Unfinished sequence: d16167P19 Contig_ID: 00879 length: 1665 bp Unfinished sequence: d16167P19 Contig_ID: 00522 length: 1285 bp Unfinished sequence: d16167P19 Contig_ID: 00672 length: 2384 bp Unfinished sequence: d16167P19 Contig_ID: 00236 length: 1961 bp Unfinished sequence: d16167P19 Contig_ID: 00871 length: 1361 bp Unfinished sequence: d16167P19 Contig_ID: 00086 length: 1445 bp Unfinished sequence: d16167P19 Contig_ID: 00243 length: 2193 bp Unfinished sequence: d16167P19 Contig_ID: 00920 length: 3441 bp Unfinished sequence: d16167P19 Contig_ID: 00962 length: 1902 bp Unfinished sequence: d16167P19 Contig_ID: 00186 length: 1815 bp Unfinished


```

sequence: djl67P19 Contig_ID: 00987 Length: 862 bp Unfinished
sequence: djl67P19 Contig_ID: 00984 Length: 1076 bp Unfinished
sequence: djl67P19 Contig_ID: 00502 Length: 1619 bp Unfinished
sequence: djl67P19 Contig_ID: 00679 Length: 4041 bp Unfinished
sequence: djl67P19 Contig_ID: 00982 Length: 4726 bp Unfinished
sequence: djl67P19 Contig_ID: 00989 Length: 843 bp.
***
*** WARNING: Phase 1 High Throughput Genome Sequence ***
***
* This sequence is unfinished. When sequencing is complete,
* the sequence data presented in this record will be replaced
* by a single finished sequence with the same accession number.
Location/Qualifiers
source
1..125492
/organism="Homo sapiens"
/clone="167P19"
/chromosome="X"

BASE COUNT 27173 a 17764 c 18531 g 29152 t 32872 others
ORIGIN
Very Match 8.9%; Score 126; DB 44; Length 125492;
Best Local Similarity 81.7%; Pred. No. 9.12e-66;
Matches 188; Conservative 0; Mismatches 38; Indels 4; Gaps 4;

Db 121496 aggaagtcgagcagcgtctggtgcaacatgcaacccgctctac-aaa-aaaaata 121553
CP 1287 AGGAGTTTGAGACGACGCTGTACACATGGCGAATGCCATCTTACCAACAAACAAACA 1228
Db 121554 gaaataatgacagaggtgtggtgcatgtgcctgtagtcccgactaccagagagctgagc 121613
CP 1227 AAAAAGTTAATCAGAGTGTGTGGTGGCCCTTCTTGTAGTTCACAGCTACT-AGGAGACTGAGG 1169
Db 121614 tggagagatcactatgagc-ccaagaggtcgagagctgagtgagtgagtgagtgagcag 121672
CP 1168 CAGAGAGATCAGCTTGAGCTCCAGAGGTCAGAGCTGCAGAGAGCCAGTGAGTGCTCCACTG 1109
Db 121673 cactccagcttgagtgacagagcagagcctgtctcaaaaaataa 121722
CP 1108 CACTCCAGCCTGGGTGACAGAGTGAAGTACCTGTCAAAAAA 1059

RESULT 9
ID HSAB1901 standard; DNA; HMW; 8339 BP.
AC AB001901;
NI d1112091
DT 26-JUL-1997 (Rel. 52, Created)
DT 26-JUL-1997 (Rel. 52, Last updated, Version 1)
DE Homo sapiens PACE4 gene, exon 4-7.
DE alternative splicing; PACE4.
OC Homo sapiens (human)
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
OC Homo.
RN [1]
RP 1-8339
RA Matsuda Y.;
RT Submitted (14-MAR-1997) to the EMBL/GenBank/DBJ databases.
RL Yoshiko Matsuda, The University of Tokushima, Department of
RL Biological Science and Technology; Minamijosanjima-cho 2-1,
RL Tokushima, Tokushima 770, Japan
RL (E-mail:matsuda@bio.tokushima-u.ac.jp, Tel:0886-56-7523,
RL Fax:0886-55-3161)
RN [2]
RA Tsuji A., Hine C., Tamai Y., Yonemoto K., Mori K., Yoshida S.,
RA Bando M., Mori K., Akamatsu T., Matsuda Y.;
RT Genomic organization and alternative splicing of human PACE4
RT (SPC4), kexin-like processing endoprotease";
RL J. Biochem. 122:438-452(1997).
FH key Location/Qualifiers
FH source 1..8339
FH /organism="Homo sapiens"

```

```

FT /sequenced_mol="DNA"
FT exon 800..910
FT /gene="PACE4"
FT /number=4
FT exon 1434..1577
FT /gene="PACE4"
FT /number=5
FT exon 2833..2909
FT /gene="PACE4"
FT exon 4912..5000
FT /gene="PACE4"
FT /number=7
SQ Sequence 8339 BP; 1979 A; 2062 C; 2033 G; 2240 T; 25 other;

Query Match 8.8%; Score 124; DB 8; Length 8339;
Best Local Similarity 79.6%; Pred. No. 2.60e-64;
Matches 187; Conservative 0; Mismatches 45; Indels 3; Gaps 3;

Db 261 ttttttttttttgaaaggctctgtgtgttgcagagactggaatgagtgagcaccat 320
QY 1062 TTTTGTGTTTGTGACAGGCTCTACACTGTGTCACAGGCTGGAGTGCAGTGCACAC- 1120
Db 321 catggtcactgagcgtctactccag-gctcaagcaatccttcaactcagctcct 379
QY 1121 CATGGCTCTGCAAGCCTTGACCTTGAGAGCTCAAGTATCCTCTGCTCAATCTCT 1180
Db 380 gagagcttgagactacagagatgcacatcatgccagctagattttttttttt 439
QY 1181 -AGTAGCTGGAAGTACAGAGAGAGGCGCACACACACACACACTACTTTTGTGTTT 1239
Db 440 tggagaggtgggttccacatgttgcagagctgtgtcctaaactcctggagctca 494
QY 1240 GGTAAAGATGCAATTTCGTCATGTTGTACAGGCTGGTCAAACTCCTAGTTCA 1294

RESULT 10
LOCUS HUMHDC 32351 bp DNA PRI 10-JUN-1994
DEFINITION Human gene for L-histidine decarboxylase, complete cds.
ACCESSION D16583
NID 9516770
KEYWORDS L-histidine decarboxylase.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Homo sapiens DNA, clones lambda HDC[1, 2, 3 and 4].
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE
1 (sites)
Yatsunami,K., Ohtsu,H., Tsuchikawa,M., Higuchi,T., Ishibashi,K.,
Shida,A., Shima,Y., Nakagawa,S., Yamachi,K., Yamamoto,M.,
Hayashi,N., Matanabe,T. and Ichikawa,R.
Structure of the L-histidine decarboxylase gene
J Biol. Chem. 269 (2), 1554-1559 (1994)
MEDLINE 94117478
REFERENCE
2 (bases 1 to 32351)
Yatsunami,K.
JOURNAL Unpublished (1993)
REFERENCE
3 (bases 1 to 32351)
Yatsunami,K.
TITLE Direct Submission
COMMENT Submitted (02-JUL-1993) to the DDBJ/EMBL/GenBank databases. Kimio
Yatsunami, Japan Tobacco INC., Pharmaceutical Basic Research Lab.;
1-13-2 Fukuura Kanazawa-ku, Yokohama, Kanagawa 236, Japan
(Tel:045-786-7690(ex.3390), Fax:045-786-7692)
Submitted (02-JUL-1993) to DDBJ by:
Kimio Yatsunami
Dept. of Pharmaceutical
Basic Research Lab
Japan Tobacco INC.
1-13-2 Fukuura, Kanazawa-ku
Yokohama, Kanagawa 236
Japan
Phone: 045-786-7690 x3390

```

FEATURES Fax: 045-786-7692.
Location/Qualifiers
source 1..32351
/organism="Homo sapiens"
/clone_11b="EMBL3"
protein_bind
/bound_moiety="GATA binding protein"
5447..5452
5460..5464
/bound_moiety="LBP-1 (leader-binding protein-1)"
5592..5597
misc_signal
/note="CAC box"
5733..5738
/note="CAC box"
misc_signal
/note="CAC box"
6124..6129
/note="CAC box"
6551..6555
/bound_moiety="LBP-1"
6788..6793
/bound_moiety="GATA binding protein"
6826..6831
/bound_moiety="GATA binding protein"
6895..6900
/bound_moiety="GATA binding protein"
6965..6969
/bound_moiety="LBP-1"
7169..7173
/bound_moiety="LBP-1"
7183..7188
/note="CAC box"
7221..7226
7252..7259
/note="TATA-like sequence"
7268..7272
/bound_moiety="LBP-1"
7278..7414
/codon_start=1
/number=1
7292..7296
/bound_moiety="LBP-1"
join(7384..7414,9600..9772,14491..14604,15461..15583,
18247..18381,18638..18781,19245..19311,20137..20289,
20391..20481,24567..24665,29671..29772,29907..30653)
/EC_number="4.1.1.22"
/codon_start=1
/product="L-histidine decarboxylase"
/db_xref="PID:d1004531"
/db_xref="PID:g516771"
/translation="MMPEEYRERGRMVDYICQYLSVREHRTPDVOPGYLAOLP
ESAPEDPDSDSIFGIERIIMPVVHWPSPHAYPALISPSLIDADALINCL
GFWASPPACTELEMVMDLAKMLGPEHFLHHHSSOGGVLOSSTSLIALLA
ARKNKILEKTSERPDDESCLNRLVAYASDOHSSVERKGLISLVKMKFLPVDDNF
LRGELQKALIEBKQKGLVFCATLGTGVCALFCLSLGPICARBGLMFLIDAY
AGTALCPERFRLKGLIEVADSTFNPSSKMMVHPCOTGFVWVKYKLOLTFSEVNI
LRHANGVATDFMHOIPLSRPVSVMVIVISFGVKNLOAHVRGTEMAKFEELV
RNDSPFEPAKRLGLVFRKLGKNCLETENLKEIKAGRLFLIPATIDOKLIRTV
TSQFTRDILRDMNLIRDAATLILSOHCSPSPVGNLISIRBARMACTSIQS
VSGAGDPVQARKITIKOPRVGKGPVKRENGLETLDPVDCFSSEAPDATKHLIS
SELESYLVQTKRKTVSLSCNSVPASAKPLTEASVKNKGSSRVRLISRPEDMM
LKSAFKKLIRKFSVSPFECSCGLQJLPCPLQAMV"
Intron
/number=1
9600..9772
/codon_start=1
/number=2
9773..14490
/number=2
14491..14604
/codon_start=1
/number=3
14605..15460
/number=3
15461..15583
/codon_start=1
CDS
protein_bind
exon
protein_bind
misc_signal
GC_signal
TATA_signal
protein_bind
exon
protein_bind
7292..7296
/bound_moiety="LBP-1"
join(7384..7414,9600..9772,14491..14604,15461..15583,
18247..18381,18638..18781,19245..19311,20137..20289,
20391..20481,24567..24665,29671..29772,29907..30653)
/EC_number="4.1.1.22"
/codon_start=1
/product="L-histidine decarboxylase"
/db_xref="PID:d1004531"
/db_xref="PID:g516771"
/translation="MMPEEYRERGRMVDYICQYLSVREHRTPDVOPGYLAOLP
ESAPEDPDSDSIFGIERIIMPVVHWPSPHAYPALISPSLIDADALINCL
GFWASPPACTELEMVMDLAKMLGPEHFLHHHSSOGGVLOSSTSLIALLA
ARKNKILEKTSERPDDESCLNRLVAYASDOHSSVERKGLISLVKMKFLPVDDNF
LRGELQKALIEBKQKGLVFCATLGTGVCALFCLSLGPICARBGLMFLIDAY
AGTALCPERFRLKGLIEVADSTFNPSSKMMVHPCOTGFVWVKYKLOLTFSEVNI
LRHANGVATDFMHOIPLSRPVSVMVIVISFGVKNLOAHVRGTEMAKFEELV
RNDSPFEPAKRLGLVFRKLGKNCLETENLKEIKAGRLFLIPATIDOKLIRTV
TSQFTRDILRDMNLIRDAATLILSOHCSPSPVGNLISIRBARMACTSIQS
VSGAGDPVQARKITIKOPRVGKGPVKRENGLETLDPVDCFSSEAPDATKHLIS
SELESYLVQTKRKTVSLSCNSVPASAKPLTEASVKNKGSSRVRLISRPEDMM
LKSAFKKLIRKFSVSPFECSCGLQJLPCPLQAMV"
Intron
/number=1
9600..9772
/codon_start=1
/number=2
9773..14490
/number=2
14491..14604
/codon_start=1
/number=3
14605..15460
/number=3
15461..15583
/codon_start=1

Query Match 8.88; Score 125; DB 99; Length 32351;
Best Local Similarity 75.7%; Pred. No. 4.87e-65;
Matches 206; Conservative 0; Mismatches 63; Indels 3; Gaps 3;
BASE COUNT 8594 a 7026 c 7405 g 9326 t
ORIGIN
Db 17093 accattatattttttaaaacaggcctgttctgtctcccaaggctgagtgagtg 17152
QY 1054 ACCACTTCTTTTCTTTTCTTTTGAAGGGTCTCACTCTGTCACCCAGGCTGGAATGCAGTG 1113
Db 17153 gaagcatcatgctcctcagcctccatcctcctg-gcccaagcaatcctccaacctca 17211
QY 1114 GCACACACAGGCTCTCTGAGCCTTGACCTTGAGGAGCTCAAGTATCTCTGCTCA 1173
Db 17212 gctcccaagctagctgagctacaggctgacacactacacactgcttttggttttgt 17271
QY 1174 GCTCCCTA-GTACTGGAAGTCAAGAAAGGCGCACACACTGACTAAGTCTTTTGT 1232
Db 17272 tttttttg-taaggaacaggctcactatgtgtccaggtgttccaactcctgagct 17330
QY 1233 TTTGTTTGGTAAGATGGCATTTTCGCATGTTGTACAGGCTGCTCAAACTCTAGGT 1292
Db 17331 caaggatcctcccaagctcgtgagttacag 17362
QY 1293 CACTTGGCTCCCAAGTGTGGAATTACG 1324
RESULT 11 HSAC000012 100000 bp DNA HTG 03-OCT-1996
LOCUS *** SEQUENCING IN PROGRESS *** EPM1/APECED region of chromosome 21,
DEFINITION BAC clone B17J3; HTGS phase 1, 13 unordered pieces.

ACCESSION	AC000012
NID	g1574906
KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 100000)
AUTHORS	Quackenbush,J., Fan,J.,-B., Cox,D.R. and Myers,R.M.
TITLE	Continuous genomic Sequence Spanning a Portion of the EPM1/ABCE1 Region of Human Chromosome 21
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 100000)
AUTHORS	Quackenbush,J., Fan,J.,-B., Cox,D.R., Myers,R.M. and Vollrath,D.
TITLE	Direct Submission
JOURNAL	Submitted (01-OCT-1996) Stanford Human Genome Center, 855 California Ave., Palo Alto, CA 94304, USA
REFERENCE	3 (bases 1 to 100000)
AUTHORS	Quackenbush,J., Fan,J.,-B., Cox,D.R., Myers,R.M. and Vollrath,D.
TITLE	Direct Submission
JOURNAL	Submitted (03-OCT-1996) Stanford Human Genome Center, 855 California Ave., Palo Alto, CA 94304, USA
COMMENT	Contact: R.M. Myers

```

*** WARNING: Phase 1 High Throughput Genome Sequence ***
***
** This sequence is unfinished. It consists of 13 contigs for
** which the order is not known; their order in this record is
** arbitrary. In some cases, the exact lengths of the gaps
** between the contigs are also unknown; these gaps are presented
** as runs of N as a convenience only. When sequencing is complete
** the sequence data presented in this record will be replaced
** by a single finished sequence with the same accession number.
***
1 2133: contig of 2133 bp in length
2134: gap of unknown length
4343 7249: contig of 2907 bp in length
7250 9459: gap of unknown length
9459 12185: contig of 2727 bp in length
12186 14395: gap of unknown length
14395 17604: contig of 3209 bp in length
17604 19812: gap of unknown length
19813 23481: contig of 3669 bp in length
23482 25382: gap of unknown length
25382 29289: contig of 3599 bp in length
29290 31499: gap of unknown length
31499 37452: contig of 5954 bp in length
37453 39660: gap of unknown length
39661 44856: contig of 5196 bp in length
44857 47064: gap of unknown length
47065 51200: contig of 4136 bp in length
51201 53408: gap of unknown length
53409 57673: contig of 4265 bp in length
57674 59881: gap of unknown length
59882 67457: contig of 7576 bp in length
67458 69665: gap of unknown length
69666 81215: contig of 11550 bp in length
81216 83423: gap of unknown length
83424 100000: contig of 16577 bp in length.

```

FEATURES

```
/organism="Homo sapiens"  
/clone="B17313"
```

```
/clone_lib-"Caltech/Research Genetics Whole-Genome BAC
library"
```

/note="Sequence data collected from 3kb subclones
(vector=pJF5 (J. -B. Fan, Stanford Human Genome Center))"

BASE COUNT	16988	a	18331	c	19442	g	18704	t	26535	others
------------	-------	---	-------	---	-------	---	-------	---	-------	--------

Query Match	8.8%	Score 125	DB 45	Length 100000
Best Local Similarity	77.4%	Pred. No. 4.87e-65		
Matches 219	Conservative 0	Mismatches 58	Indels 6	Gaps 5
Db 93847	gcccgcgtgtgtgtgctcatgcctataatcccaagactttggagccgagtgactca	93906		
Cp 1346	gccccggccctggcgaatttatgtcttatccacagacttttgggagcccaagtgaactta	1287		
Db 93907	gaagtttagagacagccctgggcaacatgtgtaaaccccgtctctat-aaa-aaatacaa	93964		
Cp 1286	ggaggttgagacacagcctgtgtaacaaacatggcgcaaatggccatcttttaccaaaaaaacaa	1227		
Db 93965	aaa--tttagccagcagatgtgtgtgtgcagcctctagtlaccagctacttvggagcttagat	94022		
Cp 1236	aaaagttagctcagcgtgtggtggcccttcccttgatgttccacttactaag-agaactgaagcc	1168		
Db 94023	ggggaagatcacctctggc-cccgaagttgaaagtgttcagatgtgttgagatcacccacatgc	94081		
Cp 1167	aggagagatcacttgtagctccacagaagctcaaaagctgcaagaagccatggtgtggccactgac	1108		
Db 94082	accgcagcctggtgtacagagcgaagacccttctcaaaaaaaaaa	94124		
Cp 1107	actccagccttggtgtacagagtgaaacacctgtgtcaaaaaaaaaa	1065		

RESULT	12
ID	HSAC12
	standard; DNA; HTG; 100000 BP.

AC	AC000012;
NI	g1574906

DT	08-OCT-1996	(Rel. 49, Created)
DT	25-JUL-1997	(Rel. 52, Last upd)

DE *** SEQUENCING IN PROGRESS *** EPM1/APECD region of chromosome 21
DE RAC clone B173L3: HMGs phase 1 13 unordered pieces

KW HTG; HTGS_PHASE1.
HTGS_PHASE1
HTGS_PHASE1

US homo sapiens (human)
OC Eukaryotae; mitochondria

OC Vertebrata; Mammalia; Primates; Catarrhini; Hominoidea;
OC Homo.

RN	[1]
RP	1-1

RA Quackenbush J., Fan J.B., Cox D.R., Myers R.M.;
RT "Contiguous Genomic Sequence Spanning a Portion

RT Region of Human Chromosome 21^a;
 BL Implicated

NAME	DATE
RN	[2]
RD	1-1-000000

RA Quackenbush J., Fan J.B., Cox D.R., Myers R.M., Vollrath D.,

Submitted (01-OCT-1996) to the EMBL/GenBank/DBJ databases.

RL Stanford Human Genome Center, 855 California Ave., Palo Alto, CA
RL 94304, USA

CC Contact: R.M. Myers Stanford Human Genome Center 855 California Ave Palo Alto, CA 94304 USA *** *** WARNING: phase CC

```

      Throughput Sequence *** ** * This sequence is
      unaffected if candidate of 12 contains four * which the error is not

```

CC known; their order in this record is * arbitrary. In some cases,

CC the exact lengths of the gaps * between the contigs are also
CC unknown; these gaps are presented * as runs of N as a convenience

cc only. When sequencing is complete, * the sequence data presented in
cc this record will be replaced *by a single finished sequence with

CC the same accession number. * 1 2133: contig of 2133 bp in length *
CC 2134 4342: gap of unknown length * 4343 7249: contig of 2907 bp in
CC

CC length * 7250 9458: gap of unknown length * 9459 12185: contig of
CC 2727 bp in length * 12186 14394: gap of unknown length * 14395

CC	length	contig	bp	in	length	* 17604	19812:	gap	of	unknown	
CC	17603:	contig	3209	bp	in	length	* 17604	19812:	gap	of	unknown
CC	length	* 19813	33481:	contig	3660	bp	in	length	* 33482	35600:	

CC gap of unknown length * 25691 29289: contig of 3599 bp in length *
CC gap of unknown length * 21400 27470: contig of 6074 bp in length *
CC length 15013 23401: contig of 3005 bp in length * 23402 23530:
CC length 15013 23401: contig of 3005 bp in length * 23402 23530:

```
CC 29290 31498: gap of unknown length * 31499 3/452: config of 3934 b
CC in length * 37453 39660: gap of unknown length * 39661 44856:
CC
```

```
CC contig01 5196 bp in Length * 44857 47064: gap of unknown Length *
CC 47065 51200: contig of 4136 bp in Length * 51201 53408: gap of
```

```
CC unknown length * 53409 57673: contig of 4265 bp in length * 57674
CC 59881: gap of unknown length * 59882 67457: contig of 7576 bp in
```

```
CC length 67458 69665 : gap of unknown length * 69666 81215: contig  
CC of 11550 bp in length * 81216 83423: gap of unknown length * 83424  
CC 100000: config of 16577 bp in length.  
FH Key Location/Qualifiers  
FT source 1..100000  
FT /organism="Homo sapiens"  
FT /clone="B173J3"  
FT /clone_1lb="Caltech/Research Genetics Whole-Genome BAC  
FT library"  
FT /note=""Sequence data collected from 3kb subclones  
FT (vector-pdF5 (J.-B. Fan, Stanford Human Genome Center))  
FT using a transposon-mediated approach."  
FT Sequence 100000 BP; 16588 A; 18331 C; 19442 G; 16704 T; 2635 other;  
SQ  
  
Query Match 8.8%; Score 125; DB 5; Length 100000;  
Best Local Similarity 77.4%; Pred. NO. 4,87e-65;  
Matches 219; Conservative 0; Mismatches 58; Indels 6; Gaps 5;  
  
Db 93847 ggcctgtgtgtgtgctcattccataatcccagcaactttggagcgagtgtagtca 93906  
1346 GCCCGGGCCTGGCAATTCATGTCGTAAATCCACACACTTTGGAGGCCAAGAATGAACCTA 1287  
Db 93907 gaagtgttagacacagccctgggacaacatgtgaaacccgcttcacat-aaa-aaatracaa 93964  
1286 GGAGTTTAGACACACCCTGTACAAACATGGCCAAAATGCCATGTTTACCAAACAATAAACAA 1227  
Db 93965 aaa--ttaaccaggcatgtgtgtgcacgcaccttaagtaaccagtaacttggaggagctgat 94022  
1226 AAAAGTATGTCAGAGTGCTGGTGGCCCTTCTGTGTAATTGCACACTACTAG-G-AGA-TGAGGC 1168  
Db 94023 ggaggatcaccttgagc-cccaagaagtggaggtgtgcagttgagttagatcaaccactgc 94081  
Cp 1167 AGGAGGATCACCTTGAGCTCCACAGAGGTCAAGGCTGCAGAGAGCCATGATGTCCTACTGC 1108  
Db 94082 accccagcctgtgtgacagagcagcagcccttcacaaaaaaa 94124  
Cp 1107 ACTCCAGCCTGGGTGACAGAGTGAAGACCTCTCAAAAAAAAA 1065  
  
RESULT 13  
ID HS399A18 standard; DNA; HTG: 145194 BP.  
AC Z97054;  
NI e1053522  
DT 20-JUN-1997 (Rel. 52, Created)  
DT 20-JUN-1997 (Rel. 52, Last updated, Version 1)  
DE Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone 339A18  
KW HTG: HTGS-PHASEL.  
KW Homo sapiens (human)  
OC Eukaryota; mitochondria eukaryotes; Metazoa; Chordata;  
OC Vertebrata; Eutheraia; Primates; Catarrhini; Hominiidae; Homo.  
RN 1  
RP 1-145194  
RA Yaodin M.:  
RT ;  
RL Submitted (06-NOV-1996) to the EMBL/GenBank/DBJ databases.  
RL Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB30 1SA,  
UK. E-mail enquiries: humtery@sanger.ac.uk Clone requests:  
CL clonerequests@sanger.ac.uk  
RL IMORANT: This sequence is unfinished and does not necessarily  
CC represent the correct sequence. Work on the sequence is in progress  
CC and  
CC the release of this data is based on the understanding that the  
CC sequence may change as work continues. The sequence may be  
CC contaminated  
CC with foreign sequence from E.coli, yeast, vector, phage etc.  
CC Order of segments is not known; 800 n's separate segments.  
CC Unfinished sequence: dU339A18 Contig_ID: 01495 Length: 13468 bp  
CC Unfinished sequence: dU339A18 Contig_ID: 01275 Length: 1181 bp  
CC Unfinished sequence: dU339A18 Contig_ID: 01198 Length: 1123 bp  
CC Unfinished sequence: dU339A18 Contig_ID: 01197 Length: 6108 bp  
CC Unfinished sequence: dU339A18 Contig_ID: 01258 Length: 1692 bp  
CC Unfinished sequence: dU339A18 Contig_ID: 01052 Length: 2161 bp
```

CC	Unfinished sequence: d0339a18	Contig_ID: 01494	Length: 3769 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 00286	Length: 866 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 00001	Length: 1048 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 00109	Length: 2530 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 00484	Length: 2883 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 01239	Length: 3632 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 00606	Length: 11827 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 00901	Length: 7102 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 01439	Length: 7531 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 00784	Length: 6928 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 00051	Length: 5877 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 00051	Length: 15570 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 00397	Length: 14465 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 00464	Length: 5758 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 01041	Length: 1598 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 00834	Length: 1951 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 00635	Length: 997 bp
CC	Unfinished sequence: d0339a18	Contig_ID: 01314	Length: 7100 bp
CH	Key	Location/Qualifiers	
FT	source	1..145194	
FT	/organism="Homo sapiens"		
FT	/clone="339a18"		
FT	/chromosome="X"		
SQ	Sequence 145194 BP; 35039 A; 29424 C; 28997 G; 33165 T; 18569 other;		
	Query Match	8.8%; Score 124; DB 5; Length 145194;	
	Best Local Similarity 79.3%; Pred. No. 2,60e-64;		
	Matches 199; Conservative 0; Mismatches 47; Indels 4; Gaps 3;		
Db	40551 tgaacccagcagcttcgagaccagccttggaaacatgacaaaccccatctcacaaaaa	40610	
Cp	1294 tgaacctgagagatttgagacacgacctgtacaaacatgacggaatgccatcttaccaaaaca	1235	
Db	40611 taaaaataaaaaaattgtccagcagtggtgtgacgcgcacccgttagttccagctacttggg	40670	
Cp	1234 --AAAAACAAAAAATTATGTGTCAGGTGTGTGTGCTCTCTGTGTACTTCAGCTACTAGG-	1178	
Db	40671 aggcctgagtggaagatctgtcttgagc-caaggtgaacgaatttcagtgagccatgac	40729	
Cp	1177 AGACCTGAGGAGGAGAGATCATTGAGCTCCAGAGAGTCAAGGCTCAGAGAGCAATGTG	1118	
Db	40730 atgcactgacatctcagcctcttgacacagatgagaccctgtctcaaaaaaaaaaaaaa	40789	
Cp	1117 GNGCCACTGCACTCCAGCGCTGGGTGCACAGAGTAGAGACCCCTGTCAAAAAAAG	1058	
Db	40790 agttgc 40795		
Cp	1057 TGTGTG 1052		
LOCUS	14		
DEFINITION	HS339a18 145194 bp DNA	HTG	20-JUN-1997
ACCESSION	Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone		
NID	339a18; HTGS phase 1.		
KEYWORDS	g2213437		
ORGANISM	HTG: HTGS_PHASE1.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;		
JOURNAL	Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae;		
COMMENT	1 (bases 1 to 145194)		
	Vaudin,M.		
	Direct Submission		
	Submitted (06-NOV-1996) Wellcome Trust Genome Campus, Hinxton,		
	Cambridgeshire, CB10 1SA, UK. E-mail enquiries:		
	humanyesanger.ac.uk Clone requests: clones@yesanger.ac.uk		
	IMPORTANT: This sequence is unfinished and does not necessarily		
	represent the correct sequence. Work on the sequence is in progress		
	and		

contaminated
with foreign sequence from E.coli, yeast, vector, phage etc. Order
of segments is not known; 800 n's separate segments. Unfinished
sequence: d3339A18 Contig_ID: 01495 Length: 1368 bp Unfinished
sequence: d3339A18 Contig_ID: 01275 Length: 1181 bp Unfinished
sequence: d3339A18 Contig_ID: 01198 Length: 1123 bp Unfinished
sequence: d3339A18 Contig_ID: 01197 Length: 6108 bp Unfinished
sequence: d3339A18 Contig_ID: 01258 Length: 1692 bp Unfinished
sequence: d3339A18 Contig_ID: 01052 Length: 2161 bp Unfinished
sequence: d3339A18 Contig_ID: 01494 Length: 3769 bp Unfinished
sequence: d3339A18 Contig_ID: 00286 Length: 868 bp Unfinished
sequence: d3339A18 Contig_ID: 00001 Length: 1048 bp Unfinished
sequence: d3339A18 Contig_ID: 00109 Length: 2530 bp Unfinished
sequence: d3339A18 Contig_ID: 00484 Length: 2883 bp Unfinished
sequence: d3339A18 Contig_ID: 01239 Length: 3632 bp Unfinished
sequence: d3339A18 Contig_ID: 00606 Length: 11527 bp Unfinished
sequence: d3339A18 Contig_ID: 00901 Length: 7102 bp Unfinished
sequence: d3339A18 Contig_ID: 01439 Length: 7531 bp Unfinished
sequence: d3339A18 Contig_ID: 00784 Length: 6928 bp Unfinished
sequence: d3339A18 Contig_ID: 00051 Length: 5877 bp Unfinished
sequence: d3339A18 Contig_ID: 00591 Length: 15570 bp Unfinished
sequence: d3339A18 Contig_ID: 00397 Length: 14465 bp Unfinished
sequence: d3339A18 Contig_ID: 00464 Length: 5785 bp Unfinished
sequence: d3339A18 Contig_ID: 01041 Length: 1885 bp Unfinished
sequence: d3339A18 Contig_ID: 00834 Length: 1591 bp Unfinished
sequence: d3339A18 Contig_ID: 00635 Length: 997 bp Unfinished
sequence: d3339A18 Contig_ID: 01314 Length: 7100 bp.

*** WARNING: Phase 1 High Throughput Genome Sequence ***
*** This sequence is unfinished. When sequencing is complete,
* the sequence data presented in this record will be replaced
* by a single finished sequence with the same accession number.
Location/Qualifiers

FEATURES
Source
1..145194
/organism="Homo sapiens"
/clone="339A18"
/chromosome="X"

BASE COUNT 35039 a 29424 c 28997 g 33165 t 18569 others

ORIGIN

Query Match 8.8%; Score 124; DB 128; Length 145194;
Best Local Similarity 79.3%; Pred. No. 2.60e-64;
Matches 195; Conservative 0; Mismatches 47; Indels 4; Gaps 3;

D3 40551 tgaaccagcagcttcgagccagccttggaacatgcaaacacccatctacaaaaa 40610
11294 TGAACCTAGGAGGTTGACGACCTGTACAACTGGCGAATGCCATCTTACCAACA 1235
D3 40611 taaaaataaaaaagttagccaggtgtgtgtagcgtgcacctgagttccagctactggg 40670
1234 - -AAAAACAAAAGTTAGTCAAGTGTGCTGCTTCTTCTGTTCCAGCTACTAGG- 1178
D3 40671 aggcgtgagtggtgaggaattcgttagc-ccaggtgacgaggttcgtgacatgac 40729
1177 AGACTGAGGAGGAGGATCACTTGAAGCTCCAGAGGTCAAGGCTGACAGAGCATGGTG 1118
D3 40730 atgcacatgacatccagcctgctgacagagtgagagccctctcctcaaaaaaaaaa 40789
1117 GTGCGACTGCACTCCAGCTGGTGACAGAGTGAAGACCTGTCAAAAAAAAAAAG 1058
D3 40790 agttgc 40795
1057 TGGTGC 1052

RESULT 15
LOCUS HSAC000118 96625 bp DNA PRI 31-JAN-1997
DEFINITION Human BAC clone RG072E11 from 7q21-7q22, complete sequence.
ACCESSION AC000118
NID g1809227
KEYWORDS HTG.

SOURCE human.
ORGANISM Homo sapiens
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 96625)
Mink,P., Graves,T. and Biewald,T.
The sequence of H. sapiens BAC clone RG072E11
Unpublished (1997)
2 (bases 1 to 96625)
REFERENCE
AUTHORS
TITLE
Waterston,R.
JOURNAL
Direct Submission
Submitted (31-JAN-1997)
Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63108, USA
http://genome.wustl.edu/gsc
e-mail: saplens@wustl.wustl.edu

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
sections once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by
sequence from more than one subclone; and the assembly was
confirmed by restriction digest.

SOURCE INFORMATION:
This clone is from the first release of the human BAC library. The
library contains cloned DNA from a human male fibroblast cell line
978SK. For references see: Shizuya et al., Proc. Natl. Acad. Sci.
89:8794-8797 (1992); Kim et al., Genomics 34:213-218 (1996).
VECTOR: pBELO
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is H_RG161K23, 200 bp overlap.
Actual start of this clone is at base position 129619 of
H_RG161K23; actual end is at 96625 of H_RG072E11

This clone contains the STS SW551727.

FEATURES
Source
Location/Qualifiers

1..96625
/organism="Homo sapiens"
/chromosome="7"
/clone="RG072E11"
/clone_1id="CTB-978SK-B"
/map="7q21-7q22"
922..994
/note="69% similarity to Human cosmid 268873
(NID:g116491)"
936..1031
/note="73% similarity to Human cosmid U15422
(NID:g642458)"
943..1056
/note="68% similarity to Human cosmid 249862
(NID:g1213360)"
1026..1090
/note="73% similarity to Human cosmid 274022
(NID:g140331)"
1109..1159
/note="76% similarity to Human cosmid 268339
(NID:g149521)"
1477..1551
/note="74% similarity to Human cosmid 273359
(NID:g1322150)"
complement(1836..2027)
/rpt_family="Alu"
2209..2268
/note="73% similarity to Human YAC U35072 (NID:g1016713)"

```

misc_feature      2366..2468
                  /note="72% similarity to Human YAC U35072 (NID:g1016713)"
repeat_region     6349..6637
                  /rpt_family="ALU"
repeat_region     6933..7243
                  /rpt_family="ALU"
misc_feature      7991..8064
                  /note="81% similarity to Human cosmid 275890
                  (NID:g1430781)"
misc_feature      7994..8059
                  /note="80% similarity to Human EST H48092 (NID:g9244144)"
repeat_region     /complement(8403..8695)
                  /rpt_family="ALU"
misc_feature      9580..9839
                  /note="100% similarity to Human EST H71176 (NID:g1042992)"
repeat_region     9840..10133
                  /rpt_family="ALU"
misc_feature      10247..10327
                  /note="72% similarity to Human cosmid 273418
                  (NID:g1322398)"
repeat_region     10509..10799
                  /rpt_family="ALU"
repeat_region     /complement(10816..10849)
                  /rpt_family="L1"
misc_feature      12410..12465
                  /note="78% similarity to Human cosmid U70984
                  (NID:g1552538)"
repeat_region     /complement(13607..13638)
                  /rpt_family="L1"
repeat_region     /complement(13847..14152)
                  /rpt_family="ALU"
misc_feature      14160..14179
                  /note="match to Human cosmid 254073 (NID:g974829)"
repeat_region     /complement(14427..14695)
                  /rpt_family="ALU"
misc_feature      15000..15055
                  /note="76% similarity to Human ABL gene U07561
                  (NID:g514262)"
repeat_region     /complement(17612..17661)
                  /rpt_family="L1"
repeat_region     /complement(17949..18216)
                  /rpt_family="ALU"
repeat_region     /complement(19222..19503)
                  /rpt_family="L1"
repeat_region     /complement(19511..19803)
                  /rpt_family="ALU"
repeat_region     /complement(20198..20908)
                  /rpt_family="L1"
misc_feature      21018..21056
                  /note="87% similarity to Human cosmid 270280
                  (NID:g1237039)"
repeat_region     22180..22445
                  /rpt_family="MER"
misc_feature      22473..22548
                  /note="75% similarity to Human cosmid 268331
                  (NID:g147582)"
misc_feature      22483..22584
                  /note="82% similarity to Human cosmid U51281
                  (NID:g1256772)"
repeat_region     /complement(23222..23515)
                  /rpt_family="ALU"
repeat_region     24106..24377
                  /rpt_family="ALU"
misc_feature      24597..24952
                  /note="match to Human EST M78483 (NID:g9272799)"
repeat_region     26387..26640
                  /rpt_family="THE"
misc_feature      26749..26862
                  /note="77% similarity to Human cosmid 275889
                  (NID:g1430780)"
misc_feature      26793..26836
                  /note="80% similarity to Human EST T51746 (NID:g653603)"
misc_feature      26808..26902

```

```

repeat_region     27108..27701
                  /rpt_family="ALU"
misc_feature      28087..28115
                  /note="89% similarity to Human PAC Z76735 (NID:g1438501)"
repeat_region     28391..28735
                  /rpt_family="MER"
misc_feature      /complement(28966..29167)
                  /note="75% similarity to Human EST H09211 (NID:g874033)"
misc_feature      /complement(29601..29652)
                  /note="94% similarity to Human EST T34443 (NID:g616541)"
misc_feature      29892..29918
                  /note="83% similarity to Human cosmid 273495
                  (NID:g1332499)"
misc_feature      30085..30121
                  /note="83% similarity to Human cosmid 273495
                  (NID:g1332499)"
misc_feature      30396..30928
                  /note="similar to Human clones U66082 (NID:g1519282) and
                  U66083 (NID:g1519284)"
misc_feature      30462..30896
                  /note="70% similarity to Human EST W93033 (NID:g1422184)"
repeat_region     /complement(31607..31884)
                  /rpt_family="ALU"
misc_feature      33037..33099
                  /note="80% similarity to Human breakpoint cluster region
                  gene U07000 (NID:g487344)"
misc_feature      33083..33205
                  /note="78% similarity to Human cosmid 268331
                  (NID:g147582)"
repeat_region     /complement(33220..33509)
                  /rpt_family="ALU"
repeat_region     /complement(34518..34755)
                  /rpt_family="ALU"
repeat_region     /complement(35928..36215)
                  /rpt_family="ALU"
misc_feature      36996..37072
                  /note="76% similarity to Human cosmid 273359
                  (NID:g1322150)"
misc_feature      37029..37131
                  /note="75% similarity to Human EST H85739 (NID:g1067318)"
repeat_region     37195..37486
                  /rpt_family="ALU"
repeat_region     37815..37851
                  /rpt_family="L1"
repeat_region     37917..38018
                  /rpt_family="L1"
repeat_region     38047..38302
                  /rpt_family="L1"

Note: remainder of annotations omitted.
Query Match      8.6%; Score 121; DB 89; Length 96625;
Best Local Similarity 78.6%; Pred. No. 3.99e-22;
Matches 180; Conservative 0; Mismatches 47; Indels 2; Gaps 2;
Db 90537 gggagttgagaccagcctgggcaacatggtgtaaacctgtctacttaaaatatgtat 90596
|||||
Db 1286 GGAGTTGAGAGACCGCTGTGACCAATGCGCAATGCGCATCTTACCAACAAAACAA 1227
|||||
Db 90597 aaactagcctgagtggtggcacatgctgtagtcctccagcctatttggagagctgggc 90656
|||||
Db 1226 AAAAGTTAGTCAGGTGGTGGCCCTTCCTTGTAGTCCAGCTACTAGG-AGACTGAGGC 1168
|||||
Db 90657 acgagaatccttgaac-caggagagcagagtgctgcagtgtagccaaagatatgacctg 90715
|||||
Db 1167 AGAGATCACTTGAGCTCCCAAGAGCAAGGCTGCAGAGAGCCATGCTGTGCTCCACTGC 1108
|||||
Db 90716 actccagcctgggtgacaaagcgagaccctgtctcaaaaaaaaaa 90764
|||||
Db 1107 ACTCCAGCCTGCTGGTGAAGAGAGTGAAGACCTGTCAAAAAAAAAAAAAA 1059
|||||

```

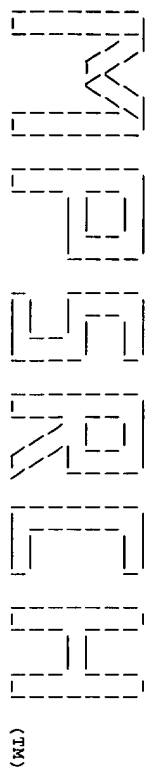
Thu Mar 26 09:16:03 1998

US-08-236-918A-7.rge

Page 13

Search completed: Wed Mar 25 13:28:43 1998
Job time : 1987 secs.

THIS PAGE BLANK (USPTO)



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligentics, Inc.

Msrch_n n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Wed Mar 25 13:54:43 1998; MasPar time 384.28 Seconds
928.964 Million cell updates/sec
Near output not generated.

Title: >US-08-236-918A-7
Description: (1-1415) from 5674704.seq
Perfect Score: 1415
N.A. Sequence: 1 AGTGGAAAGTCTCCGCGAG.....TGATAAAAA.....AAAAA 1415
Comp: TCACCTTCAAGAGCGCGTC.....ACTATTTT.....TTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 333433 seqs, 126143548 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

STS
1:STS1 2:STS2 3:STS3 4:STS4 5:STS5 6:STS6 7:STS7 8:STS8
9:STS9 10:STS10 11:STS11 12:STS12 13:STS13 14:STS14
15:STS15 16:STS16 17:STS17 18:STS18 19:STS19
20:STS20 21:STS21 22:STS22 23:STS23 24:STS24
25:STS25 26:STS26 27:STS27 28:STS28 29:STS29
30:STS30 31:STS31 32:STS32 33:STS33 34:STS34
35:STS35 36:STS36 37:STS37 38:STS38 39:STS39
40:STS40 41:STS41 42:STS42 43:STS43 44:STS44
45:STS45 46:STS46 47:STS47 48:STS48 49:STS49
50:STS50 51:STS51 52:STS52 53:STS53 54:STS54
55:STS55 56:STS56 57:STS57 58:STS58 59:STS59
60:STS60 61:STS61 62:STS62 63:STS63 64:STS64
65:STS65 66:STS66 67:STS67 68:STS68 69:STS69
70:STS70 71:STS71 72:STS72 73:STS73 74:STS74
75:STS75 76:STS76 77:STS77 78:STS78 79:STS79
80:STS80 81:STS81 82:STS82 83:STS83 84:STS84
85:STS85 86:STS86 87:STS87 88:STS88 89:STS89
90:STS90 91:STS91 92:STS92 93:STS93 94:STS94
95:STS95 96:STS96

Statistics: Mean 12.191; Variance 5.851; scale 2.084

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1415	100.0	1415	9	G28572 human STS SHGC-35316.	0.00e+00
2	300	21.2	515	8	G26790 human STS STS_U03397.	1.78e-222

3	115	8.1	373	31	AA470779	nel9e10.s1	NCI_CGAP_C	3.32e-65
4	115	8.1	428	80	HSAA70779	nel9e10.s1	NCI_CGAP_C	3.32e-65
5	115	8.1	428	80	HSAA70779	nel9e10.s1	NCI_CGAP_C	3.32e-65
6	115	8.1	428	80	HSAA70779	nel9e10.s1	NCI_CGAP_C	3.32e-65
7	114	8.1	503	52	AA488486	aa53b01.s1	NCI_CGAP_G	3.32e-65
8	114	8.1	503	52	AA488486	aa53b01.s1	NCI_CGAP_G	3.32e-65
9	114	8.1	503	52	AA488486	aa53b01.s1	NCI_CGAP_G	3.32e-65
10	111	7.8	284	37	AA482730	zt12d11.r1	NCI_CGAP_G	2.15e-64
11	111	7.8	284	37	AA482730	zt12d11.r1	NCI_CGAP_G	2.15e-64
12	109	7.7	323	66	HS1274183	nf47g04.s1	NCI_CGAP_A	5.84e-62
13	108	7.6	260	38	AA502565	nf47g04.s1	NCI_CGAP_A	5.84e-62
14	108	7.6	260	38	AA502565	nf47g04.s1	NCI_CGAP_A	5.84e-62
15	108	7.6	300	20	AA262532	ne88g02.s1	NCI_CGAP_K	2.42e-60
16	108	7.6	300	20	AA262532	ne88g02.s1	NCI_CGAP_K	2.42e-60
17	108	7.6	338	68	HS1288482	ng62f10.s1	NCI_CGAP_L	1.56e-59
18	108	7.6	338	68	HS1288482	ng62f10.s1	NCI_CGAP_L	1.56e-59
19	108	7.6	338	68	HS1288482	ng62f10.s1	NCI_CGAP_L	1.56e-59
20	108	7.6	338	68	HS1288482	ng62f10.s1	NCI_CGAP_L	1.56e-59
21	107	7.6	399	30	AA487199	zsl1c02.r1	NCI_CGAP_G	1.56e-59
22	107	7.6	399	30	AA487199	zsl1c02.r1	NCI_CGAP_G	1.56e-59
23	108	7.6	447	76	HSAA27643	zsl1c02.r1	NCI_CGAP_G	1.56e-59
24	107	7.6	460	42	AA525018	aa46a10.s1	NCI_CGAP_G	1.56e-59
25	107	7.6	460	42	AA525018	aa46a10.s1	NCI_CGAP_G	1.56e-59
26	106	7.5	328	71	HS1309371	nc76a08.r1	NCI_CGAP_P	1.56e-59
27	106	7.5	328	71	HS1309371	nc76a08.r1	NCI_CGAP_P	1.56e-59
28	106	7.5	328	71	HS1309371	nc76a08.r1	NCI_CGAP_P	1.56e-59
29	105	7.4	328	68	AA469451	nc76a08.r1	NCI_CGAP_P	1.56e-59
30	105	7.4	328	68	AA469451	nc76a08.r1	NCI_CGAP_P	1.56e-59
31	105	7.4	328	68	AA469451	nc76a08.r1	NCI_CGAP_P	1.56e-59
32	105	7.4	328	68	AA469451	nc76a08.r1	NCI_CGAP_P	1.56e-59
33	105	7.4	328	68	AA469451	nc76a08.r1	NCI_CGAP_P	1.56e-59
34	104	7.3	341	79	HSAA68319	nc77a05.s1	NCI_CGAP_P	2.59e-56
35	104	7.3	341	79	HSAA68319	nc77a05.s1	NCI_CGAP_P	2.59e-56
36	104	7.3	361	49	AA229986	nc37d05.s1	NCI_CGAP_P	2.59e-56
37	104	7.3	361	49	AA229986	nc37d05.s1	NCI_CGAP_P	2.59e-56
38	104	7.3	372	28	AA468388	nc78f10.s1	NCI_CGAP_P	2.59e-56
39	104	7.3	372	28	AA468388	nc78f10.s1	NCI_CGAP_P	2.59e-56
40	104	7.3	382	68	AA502532	ng22f09.s1	NCI_CGAP_O	2.59e-56
41	104	7.3	382	68	AA502532	ng22f09.s1	NCI_CGAP_O	2.59e-56
42	104	7.3	430	28	AA468484	nc78d10.r1	NCI_CGAP_P	2.59e-56
43	104	7.3	430	28	AA468484	nc78d10.r1	NCI_CGAP_P	2.59e-56
44	104	7.3	486	43	AA528706	nf51e11.s1	NCI_CGAP_A	2.59e-56
45	104	7.3	486	43	AA528706	nf51e11.s1	NCI_CGAP_A	2.59e-56

ALIGNMENTS

RESULT 1 G28572 1415 bp DNA STS 01-JUL-1996
LOCUS human STS SHGC-35316.
DEFINITION G28572
ACCESSION g1408387
KEYWORDS STS sequence; primer: sequence tagged site.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Euteria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1415)
AUTHORS Myers, R.M.
JOURNAL Unpublished (1996)
COMMENT

Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, W-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myerseshgc.stanford.edu
Primer A: TGAATGGAAGTCAATAGGC
Primer B: AAAAAGTGTGCAATTTTAAAGC
STS size: 180
PCR Profile:

Initial incubation: 94 degrees C for 90 seconds	
Denaturation: 94 degrees C for 15 seconds	
Annealing: 62 degrees C for 23 seconds	
Polymerization: 72 degrees C for 30 seconds	
PCR Cycles: 30	
Thermal Cycler: Perkin Elmer 9600	
Protocol:	
Template:	25 ng
Primer:	each 1 uM
dNTPs:	each 200 uM
Tag Polymerase:	0.05 units/ul
Total Vol:	10 ul
Buffer:	
MgCl2:	2.5 mM
KCl:	50 mM
Tris-HCl:	20 mM
pH:	8.3
Prepared with primer pairs provided by Sandoz, derived from U03397	
-- Washington University/Merck EST sequence.	
Location/Qualifiers	
source	1..1415
organism	"Homo sapiens"
map	1
STS	885..1064
primer_bind	885..905
primer_bind	complement(1041..1064)
BASE COUNT	385 a 332 c 333 g 365 t
ORIGIN	
Query Match	100.0%; Score 1415; DB 9; Length 1415;
Best Local Similarity	100.0%; Pred. No. 0.00e+00;
Matches	1415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	1 atgtgaagaattctccggcagcccttgatgttcaagatgtgacatttgtgaacacgccaat 60
Oy	1 AATGTGAAAGTTCTCCCGCAGCCCTGAGATCTCAAGATGACATTGTGAGACAGCTAAT 60
Db	61 ttgatataaattctctcttggaatcagcttgtagtatacataccctgtgcagatttcatca 120
Oy	61 TTGATTAAATTTCTCTTGGAATCAGCTTCTGATCATACCTGTGCGAGATTTCATCA 120
Db	121 tgggaacacgcgtgttaacaatagtagccacatctgttctgttctcaacttggagaga 180
Oy	121 TGGGAACACGCTGTTCACACATAGTAGCCACTGTGTGCTGCTCCTCAACTTTGAGAGA 180
Db	181 caagatcatgtcagagttcctttagtaactgtccacagctggttaacttctgtgataataca 240
Oy	181 CAAGATCATGTGAGAGTTCCTTGTAGTACTGCCAGCTGCTACATTTCTGATTAATTACA 240
Db	241 ggaatcagaattgcagctccctgcctcccaaatagttcttcacgcagcaggtggacaaga 300
Oy	241 GGAATCAGATTTCGCACTCCCTGCTCTCCCAATATGTTTCTCAGCGCAGGTGGACAAAGGA 300
Db	301 cctgtacacatgacagcagcagcgttaaaagtggttttcagagaccaggaagagtgctctcca 360
Oy	301 CCTGTACATATGACAGCAGCAGTGTAAAGGTGTTTTCAGAGACCAAGGAAGAGTGTCTCTCCA 360
Db	361 ccagacatgacagatgtgtgactgacatccacaggttttaactgctctggggcagagatgcaca 420
Oy	361 CCAGACATGACAGATGTGTGACTGCACATCCAGAGGTTTCACTCCCTGGGGCGAGATGCACGA 420
Db	421 tgtgtgaacagagattgtaaacaagctcaagaactgtgcaaaaaaaggttgtaaagactgttc 480
Oy	421 TGTGTGAACAGAGATTGTAAACAAGCTCAAGAACTGCAAAAATAAGGTGTGAAAGACTGTT 480
Db	481 gctttggagacatttaacagatcagaacacgttgcactctgtcgacacctgtgacaacactgttct 540
Oy	481 GCTTTGGGACATTTAACGATCAGAAACGTGGGACATCTGTGCACCCCTGACAAACCTGTTCT 540
Db	541 tggatggaagctctgtgtcttgtgaatcgggacgaagagaggaagcgtgtgtctgtgacacat 600

QY	541	TTGATGGAAATCTGTCTTGTGAATGGAGCGAAGGAGGAGGACGTGGCTGTGTGACCAT	600
Db	601	ctccagccgaacctctccccggagacatccctctgtgaccccccgctgcacctgagaaagc	660
QY	601	CTCCAGCGGAGACTCTCTCGGGAGACATCCTCTGTATACCCCGCTGCCCTGGAGAGAC	660
Db	661	caggacactctccgagatcatctctctctctgtgcgtgacgtgcagctgctctct	720
QY	661	CAGGACACTCTCCGACATCATCTCCTCTCTTCCGTGACGTGCACTGGGTGCTCT	720
Db	721	tcgcgctgtctctctacgctcgcgtttctcgtgtgtaaggggcaggaanaaaccc	780
QY	721	TCTGCTGTCTTCTCTACCCCTCCGTTTCTGTGTGTTAAAGGGGGCGAAAGAACTCC	780
Db	781	tgtatataatcaacaacccattatagaccagiatacaactactcaagaggaaagtgc	840
QY	781	TGTATATATTCAAACACACCATTTATAGACACAGTACAAACTCTCAAGAGAAAGTGCT	840
Db	841	gtacgtgcgaattcccgaaagaaaggaagatgtgtgtgtgtgtgtgtgtgtgtgtgt	900
QY	841	GTACTGTGCGGATTTCCAGAGAGAGAGAGAGAGATGTGAATGTGAATGGAATCAAT	900
Db	901	aggcgtgtgtggacctttcttgaaaagaagcaagaaatatgtgtcatccgtatcaagc	960
QY	901	AGGCGTGTGTGGACTTTCTTGAAAGAACCAAGAAATGTGATCATCGCTATCACAGC	960
Db	961	ttcaaaagcaagaacacccatccatataatccccagatcccccaacacgttctt	1020
QY	961	TTTCAAAAGCAAGACACACATCTCTACATATATACCAAGATTTCCCAACACAGCTTCTT	1020
Db	1021	tctaatgtgcaatgtgtgtgctcttaaaaaatgcacactttttttttttttttgtacag	1080
QY	1021	TCTAAATGCCAATGAGTTGGCTTTAAAAATGCACACTTTTTTTTTTTTGTGACAGG	1080
Db	1081	gttcacattcgttaacccaagctgtgaatgtcaatgtgcacacacatgtctctgcagcttg	1140
QY	1081	GTCACACTCTGTCAACCCAGGCTGTGAGTGCAGTGCACCACTGGCTCTGTGACGCTTG	1140
Db	1141	accctctggagctcaagtgtctctctgctcctcaatctccttaagctgtgaactacaaga	1200
QY	1141	ACCTCTGGGAGCTCAAGTATCTCTCTGCTCAAGTCTCTACTAGCTGAACACTACAGGA	1200
Db	1201	agggccacacacacctgtactacttttgttttgtttgtttgtttgtttgtttgtttgtt	1260
QY	1201	AGGGCCACACACCTGTACTACTTTTGTTTTGTTTGTGTTGTTGTTGTTGTTGTTGTTG	1260
Db	1261	tgttgtacagctgtgtctcaaatctcttaagttcaactttgtgcctcccaagtgctggagt	1320
QY	1261	TGTTGTACAGCTGTGTCTCAAACTCTTAGTTCACTTTGGCTCCCAAGTGTGGGATT	1320
Db	1321	acagacactgacatgcagagcccgagcaaaatgatgacacacttttaacagaagagacga	1380
QY	1321	ACAACATGAACTGCCAGGCGCCGCCAAAATATATGACACCACTTTTAAACAGACACAGA	1380
Db	1381	tgaagacagagctgtgtatataaaaaaaatgaaatgaaatgaaatgaaatgaaatgaaatg	1445
QY	1381	TGAGGACAGAGCTGTGATTAATAAAAAAATAAAAAA 1415	
RESULT	2		
LOCUS	G26790	515 bp	DNA
DEFINITION	human STS STS_U03397.		
ACCESSION	G26790		
NTD	g1349022		
KEYWORDS	STS sequence; primer: sequence tagged site.		
SOURCE	human STS derived from sequences in dbEST and the Unigene collection.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotes: mitochondrial eukaryotes; Metazoa: Chordata: Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 515) Hudson,T.		

QY	1059	TTTTTTTTTTTTTTTTTGACAGGGCTCAGCTCTGTACACAGGCTGAGAGTCAGTGGACAC	1118
Db	62	atcatggttactctcagcctctgaattctcgtg_gctcaagtgatctctcgtcctcagcct	120
OY	1119	ACCAATGCGTCTCTCAGACCCTTGACCTCTGGAGACTCAATGTAGTCCCTGCTCAGTCT	1177
Db	121	ccccatggtgtgggactcagaaagaagcaccacacccagctcga- -tttttatttttctt	179
OY	1178	CCCTAGTGCCTGGAACTCAAGAGGAAGGGCCACACACTACATCAACTTTTGTGTTTTTGT	1237
Db	180	ttt- -taaaagcagagctctcgtctatgtttatgacagcttattttgaactctcgtgtt	233
OY	1238	TTGGTAAAGATGAGCATTTGCCCAATGTTGACAGGCTGGTCTCAAACTCCTAGGTT	1292

RESULT	4
ID	HSA070779 standard; RNA; EST; 373 BP.

AC	AA470779;
NI	g2198088
DT	21-JUN-1997 (Rel. 52, last updated)
DE	28-JUN-1997 (Rel. 52, last updated, Version 3)
DE	neig9c0.s1 NCI-CGAP_C03 Homo sapiens cDNA clone 881706 similar to
DE	contains Alu repetitive element/contains element LTR10 repetitive
DE	element ;.
KM	EST.
OS	Homo sapiens (human)
OC	Eukaryotae; mitochondrionl eukaryotes; Metazoa; Chordata;
OC	Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
OC	Homo.
RN	[1]
RP	1-373
RA	NCI-CGAP;
RT	"National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
RT	Tumor Gene Index";
RL	Unpublished.
CC	Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email:
CC	Robert.Strausberg@nih.gov Tissue Procurement: Elias Campo, M.D.,
CC	Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: M.
CC	Berto Soares, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D. D.
CC	Sequencing by: Washington University Genome Sequencing Center Clo
CC	distribution: NCI-CGAP clone distribution information can be found
CC	through the I.M.A.G.E. Consortium/INM at:
CC	www.hio.llnl.gov/dbp/image/image.html Insert Length: 1322 Std
CC	Error: 0.00 Seq primer: -41m13 fwd. Ef from Amersham High quality
CC	sequence stop: 359.
EH	Key Location/Qualifiers

```

FT      source
FT      1..373
FT      /organism="Homo sapiens"
FT      /note="Vector: p773D-Pac (Pharmacia) with a modified
FT      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FT      was prepared from 12 pooled bulk tumor samples and primed
FT      with a Not I - oligo(dT) primer. Double-stranded cDNA was
FT      ligated to Eco RIadaptors (Pharmacia), digested with Not I
FT      and cloned into the Not I and Eco RI sites of the modified
FT      p7733 vector. Library went through one round of
FT      normalization."
FT      /clone="881706"
FT      /clone_1ib="NCI_CGAP_Cc3"
FT      /sex="pooled"
FT      /tissue.type="colon"
FT      /lab_host="DH10B"
FT      <1..>373
FT      mRNA
FT      Sequence 373 BP; 90 A; 89 C; 88 G; 106 T; 0 other;
FT      Q

```

Query Match	8.18; Score 115; DB 80; Length 373;
-------------	-------------------------------------

Best Local Similarity	78.78;	Pred. No. 3.32e-65;
Matches	185; Conservative	0; Mismatches 46; Indels 4; Gaps 4

Db 2 tttcttcttcttgagacagggtcactctglttaacctagctgggtgcagtgcaca 61
|||||
|||
Qy 1059 ttttttttttttttttttttcacaggggtccatctgtcacaccacaggctggactgcactggcacc 118
|||||

Db	62	atcagtgttcaacgcgcagccctgatattcccg-gctaaatgatctctcgtcctaagcctt	120
Qy	1119	ACCATGCTCTCTGCACACCTTGACCTCTGGAGACTCAATGATCCCTCGCTCAGTCT-	1177
Db	121	cccagtagtggggactacagaagaagcaccacacccagcgcga- <u>tttttat</u> ttttctt	179
Qy	1178	CCCTGCTAGCTGGAACTCAAGGAAGGCGCACCCACACTGACATCACTTTTCTTTTGT	1237
Db	180	ttg-taaagacagagctctctgatagttgaatgaagcctatttgaactcctggct	233
Qy	1238	TTGGTAAAGAGGCGATTTCCGCATGTTGTAACAGCGCTGATCAAACTCCTAGGTT	1292

RESULT	5
ID	HS1285256 standard; RNA; EST; 428 BP.

AC AA488846;
NI g2218448
DT 01-JUL-1997 (Rel. 52, Created)
DT 01-JUL-1997 (Rel. 52, last updated, Version 1)
DE aa55b01.st NCI_CGAP_GCB1 Homo sapiens cDNA clone 824809 3' similar
DE to contans Alu repetitive element'.
KW EST.
RV
OS Homo sapiens (human)
OC Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
OC Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
OC Homo.
RN (1)
RP 1-428
RA NCI-CGAP;
RT "National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
RT Tumor Gene Index";
RL Unpublished.
CC Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email:
CC Robert.Strausberg@nih.gov Tissue Procurement: Louis M. Staudt
CC M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D. cDNA library
CC Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. cDNA
CC Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by:
CC Washington University Genome Sequencing Center Clone distribution:
CC NCI-CGAP clone distribution information can be found through the
CC M.A.G.E. Consortium/ILIN. at:
CC www-bio.linnl.gov/borp/image/image.html Seq primer: -41m3 fwd. ET
CC from Amersham High quality sequence stop: 418.

	Location/Qualifiers
FH Key	
FH	
FT	1..428
source	

FT	/organism="Homo sapiens"
FT	/note="Vector: pT73D-Pac (Pharmacia) with a modified
FT	polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand
FT	was prepared from human tonsillar cells enriched for
FT	germinal center B cells by flow sorting (CD20 ⁺ , IgD ⁻),
FT	provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
FT	(NCI), and Dr. Gerald Marti (CGBR). cDNA synthesis was
FT	primed with a Not I - oligo(dT) primer
FT	[5'-TGTTCACCAATCGTAGTGAGCGGCCGCTCAATTTTTTTTTTTTTT-
FT	3']. Double-stranded cDNA was ligated to Eco RI adaptors (
FT	Pharmacia), digested with Not I and cloned into the Not I
FT	and Eco RI sites of the modified pT73 vector. Library was
FT	+
FT	through one round of normalization, and was constructed by
FT	Bento Soares and M. Fatima Bonaldo."
FT	/clone="824809"
FT	/clone_id="NCI-CGAP-GCBI"
FT	/tissue_type="germinal center B cell"
FT	/lab_host="DH10B"
FT	complement(<1..>428)
FT	mRNA

Sequence 428 BP; 109 A; 106 C; 91 G; 122 T; 0 other;

Query Match 8.18; Score 115; DB 67; Length 428;

Matches 187; Conservative 0; Mismatches 48; Indels 4; Gaps

Ddb 2 ttttttttttttggaggggtctcaactcttgacaccgaactgtagtcagtgtgc 61
|||||
|||||

Query Match	8.1%	Score 115	DB 30	Length 428
ORIGIN	109 a	106 c	91 g	122 t
BASE COUNT	mRNA complement(<1..>428) /lab host="DH10B" /clone="824809" /clone_lib="NCI_CGAP_GCB1" /tissue_type="germinal center B cell" /lab host="DH10B"			
Accession	AA488846 428 bp mRNA EST 25-JUN-1997			
Definition	aa55b01.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone 824809 3' similar to crotalus Alu repetitive element;.			
Accession	AA488846 92218448			
Source	EST.			
Organism	human.			
Source	Homo sapiens			
Source	Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrate; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;			
Reference	Homo. 1 (bases 1 to 428)			
Authors	NCI-CGAP.			
Title	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index			
Journal	Unpublished (1997)			
Comment	Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.Strausberg@nih.gov Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D. cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: www-bio.llnl.gov/bbrp/image/image.html Seq primer: -41m3 fwd. ET from Amersham High quality sequence stop: 418. Location/Qualifiers 1..428 /organism="Homo sapiens" /note="vector: pRT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCGCCCTCATTTTCTTTT-3'] Double stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."			

	D	b	Best Local Similarity 78.2%; Pred.NO. 3,326-65;
Matches	187; Conservative	0; Mismatches	48; Indels 4; Gaps 4;
FT	F	T	
TT	t	t	
TTTT			
QY	1059 TTTTGTGCTGGAGCGTGTCACACTC-TGCACCACGAGCGATCCAGCGCAC	1117	
Dd	62 ataccacgtcactgcttgcttcgcgttccatcttgtcacccagcaagtgaagtagttgac	120	
Oy	1118 CACCATGGCTCTTCGACACCTTGGACCTTGGAGGTCAAGTAGTCTCTGCTAGTCT	1177	
Dd	121 ccctgatggcgaggacctaacgcatlgygccaccaacacggycacaattatlttgatttic	180	
Oy	1178 CCT-AGTAGCTGAACACTACAAGAAG-GGCCAACCCACACCTGCATTAACTTTTGT	1235	
Dd	181 ttttttcagaagcgggttcacccaatgttgcceagagctgtgtcttaactcctgtgcta	239	
Oy	1236 GTTTGTAAAATGGCATTTTGC CATGTGTACAGGGCGTGCTCAAACCTTAGGTTC A	1294	
RESULT	7 standard; RNA; EST; 503 BP.		
ID	HSL186885		
AI	AZ828360;		
NI	G1925276		
DT	04-Apr-1997 (Rel. 51, Created)		
DD	20-JUN-1997 (Rel. 52, Last updated, Version 3)		
DE	Zrl2dl1.rl NC1_GCAP_GCBI Homo sapiens cDNA clone 712917 5' similar to contains Alu repetitive element;contains element PTR7 repetitive element ; . DE EEST. KW		
OS	Homo sapiens (human) OC Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; NC Vertebrates; Eutheria; Primates; Catarrhini; Homnidae; Homo. RN [] RP I-503 RA NC1-GCAP; RT "National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index"; RL Unpublished. CC Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.Strausberg@nih.gov This clone is available royally-free through LNLN ; contact the IMAGE Consortium (infoimage.llnl.gov) for further information. Seq primer: -28ml3 rev2 ET from Amershams High quality sequence stop: 483. FH Key Location/Qualifiers FI source		
FT	source		
FT	1..503		
FT	/organism="Homo sapiens"		
FT	/note="Vector: pYT713D-PAC (Pharmacia) with a modified		
FT	polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand CDNA		
FT	was prepared from human tonsillar cells enriched for		
FT	germinal center B cells by flow sorting (CD20+, IgD+),		
FT	provided by Dr. Louis M.Staudel (NCI). Dr. David Allman		
FT	(NCI) and Dr. Gerald Marti (CSBR). cDNA synthesis was		
FT	primed with a Not I - oligo(dt) primer		
FT	[5'-GTGTACCAATCTGAAGTGGAGCGGCCGCTCATTTTTTTTTTTTTTT-		
FT	3']. Double-stranded cDNA was ligated to Eco RI adaptors (the		
FT	Pharmacia), digested with Not I and cloned into the Not I		
FT	and Eco RI sites of the modified pYT713 vector. Library wem		
FT	t		
FT	through one round of normalization, and was constructed by		
FT	Bento Soares and M. Fatima Bonaldo."		
FT	/clone-"712917"		
FT	/clone.lib="NC1_GCAP_GCBI"		
FT	/tissue_type="germinal center B cell"		
FT	/lab_host="DH10B"		
FT	<1..>503		
SQ	mRNA		
	Sequence 503 BP; 138 A; 114 C; 146 G; 105 T; 0 other;		
	Query Match	8.1%; Score 114; DB 52; Length 503;	
	Best Local Similarity 76.3%; Pred.No. 2,156-64;		
Matches	209; Conservative	0; Mismatches	59; Indels 6; Gaps 5;

Db	77	gtcaacgccttataaccacgacacttctggcggatcaactttagcccaagattttgagaca	136
Cp	1332	gtttatgtctgtatattccacgacctttggagaccCAAGTGAACCTTAGGAGTTGAGACCA	1273
Db	137	accttgggaagaatgycgagagaccacatctctac--aa-aaaaatttaaaa-ttacctggg	192
Cp	1272	gccctttacCAACATGCGAGAAATGCCATCTTTACCAACAAAAAACAATAAGTTACTCAGG	1213
Db	193	catgtttgcacacacactttatcttcaagcttaccgaagtctgaagttgggaagatcccttg	252
Cp	1212	ttgtgtgtgccccttctctttagtttccagctactc-AGGAGctTAGGcAGAGGATCACTTG	1154
Db	253	agc--ccaggtgtcaagcgtctgtgtgactatgatcatgacatgcacactcaccagccttggt	311
Cp	1153	AGCTCCAGAGGTCAAGGCTGCAGAGAGGCCAATGtGTGTGCCACTGCACCTCGTGGT	1094
Db	312	agcgaacaagaacccctgtctccaataaaaaaaa	345
Cp	1093	GACAGAGTGAGACCTGTCAAAAAAAAAAAAAA	1060

FOCUS	AA282360	503 bp	mRNA	EST	13-JUN-1997
DEFINITION	z12d11.t1	NCI_CGAP_CG41	Homo sapiens	CDNA clone 712917.5'	similar to contigins Alu repetitive element; contains element F1R7 repetitive element ;
ACCESSION	AA282360				
NID	g1925276				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrates; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 503)				
AUTHORS	NCI_CGAP.				
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT					

Contact: Robert Strausberg, Ph.D.
Tel.: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 rev2 Et from Amer sham
High quality sequence start: 483.

FEATURES

1..503

```

/note/Vector: pRT17-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I. Site_2: Eco RI. 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with the Not I - oligo(dT) primer
15'-TGTTACCATCTGTAATGGAGACGGCCGCTCATTTTTCCTTTTTC-
3'. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pRT3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo.
/cclone_1=712917
/cclone_1ld="NCI_GCAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"

```

MRNA	<1..>503		
BASE COUNT	138 a	114 c	146 g
ORIGIN			105 t

Query Match	8.18;	Score 114;	DB 22;	Length 503;
Best Local Similarity	76.38;	Pred. No. 2.15e-64;		

Matches	209;	Conservative	0;	Mismatches	59;	Indels	6;	Gaps	5;
---------	------	--------------	----	------------	-----	--------	----	------	----

Db	77	gttaagccctgatatccccagcaatgctggtggcgagatcactgagccagggaattgagacca	136
Cp	1332	gtttatgtctgtatattccacacctgttggagagccaaatgaaactttagagtttgagacca	1273
Db	137	accctggcaagatggtgcgagaccccatctctac - aa - aaaaattaaaaa - ttaactggg	192
Cp	1272	gcctgtgtacaaataggagaaatgccatctttagccaaacaaaaacaaaaaaagttatgtacgg	1213
Db	193	catgtgtgcacacacacttatcttcaagctactcaagctgtgctgaggttggagagatccctt	252
Cp	1212	tgtgtgtgccccttccctttagtgcacgacgt - agagagcttagggcaagagatcacttg	1154
Db	253	agc - ccagctggtcacaaggtcgtgtgtaagtcagatcatgctacgtcactcgaagcctggt	311
Cp	1153	agctcccaagaggtcaaggctgcagaaagccatggtgtgtccctcactcatccagcctgggt	1094
Db	312	agcagacaagaagccctgcttcaaaaaaaaaaaaaa	345
Cp	1093	gacagagtgagacctgtctcaaaaaaaaaaaaaa	1060

RESULT	9	AA482730	284 bp	mRNA	EST	08-JUL-1997
LOCUS		n147904.s1	NCI-CGAP_Alvi	Homo sapiens	CDNA clone	923094 similar to
DEFINITION		contains Alu repetitive element;.				
ACCESSION		AA482730				
NID		92211575				
KEYWORDS		EST.				
SOURCE		human.				
ORGANISM		Homo sapiens				
		Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;				
		Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae;				
		Homo.				
REFERENCE		1 (bases 1 to 284)				
AUTHORS		NCI-CGAP.				
TITLE		National Cancer Institute, Cancer Genome Anatomy Project (CGAP),				
JOURNAL		Tumor Gene Index				
		Unpublished (1997)				

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.11nl.gov/dbirp/image/image.html

FEATURES

Source

```

/organism="Homo sapiens"
/note="Vector: pMFP10; mRNA made from al
rhabdomyosarcoma, cDNA made by oligo-dT
Non-directionally cloned. Size-selected
average insert size 600 bp. Reference:
1996) Cancer Research 56:5580-5583."
/clone="923094"
/clone_id="NCL CGAP Alvi"
/tissue_type="alveolar rhabdomyosarcoma"
/lab_host="DH10B"

```

ORIGIN	BASE COUNT	MRNA
97 a	71 c	<1..>284
	72 g	
	44 t	

Query Match	7.8%;	Score 111;	DB 37;	Length 284;
Best Local Similarity	78.3%;	Pred. No. 5.84e-62;		


```

AC AA502565;
NI 92237532
DT 04-JUL-1997 (Rel. 52, Created)
DT 12-JUL-1997 (Rel. 52, Last updated, Version 2)
DE n962f10.s1 NCI-CGAP_Lip2 Homo sapiens CDNA clone 939403 similar to
DE contains Alu repetitive element;.
KW EST.
OS Homo sapiens (human)
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae;
OC Homo.
RN [1]
RP 1-260
RA NCI-CGAP;
RT "National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
RT Tumor Gene Index";
RT Unpublished.
CC Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email:
CC Robert.Strausberg@nih.gov Tissue Procurement: L. Jeffrey Medeiros;
CC M.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation:
CC David B. Kitzman, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D.
CC DNA Sequencing by: Washington University Genome Sequencing Center
CC Clone distribution: NCI-CGAP clone distribution information can be
CC found through the I.M.A.G.E. Consortium/LINM at:
CC www.bio.lnlni.gov/dbrrp/image/image.html Insert Length: 561 Scd
CC Error: 0.00 Seq primer: -40m13 fwd. ET from Amersham High quality
CC sequence stop: 133.
FH Key Location/Qualifiers
FT source
FT 1..260
FT /organism="Homo sapiens"
FT /note="Vector: pMP10; mRNA made from liposarcoma, CDNA
FT made by oligo-dt priming. Non-directionally cloned.
FT Size-selected on agarose gel, average insert size 600 bp.
FT Reference: Kitzman et al. (1996) Cancer Research
FT 56:5380-5383."
FT /clone="939403"
FT /issue_1lb="NCI-CGAP_Lip2"
FT /issue_type="liposarcoma"
FT /lab_host="DH10B"
FT mrna
FT <1..>260
SQ Sequence 260 BP; 47 A; 73 C; 60 G; 80 T; 0 other;

Query Match 7.6%; Score 108; DB 69; Length 260;
Best Local Similarity 79.8%; Pred. No. 1.56e-59;
Matches 178; Conservative 0; Mismatches 40; Indels 5; Gaps 5;

2 ttttttttttttttgagacaggggtcttcgcctgtc-cccagcctgagtgatgcatg 60
|||||
1059 ttttttttttttttttgacagggctctcactctgtcaccacagcctgagtgacgacac 118
|||||

Db 61 gtacagtcctcagcctcagcctcaacctccag-gcccaagtgatcctcagcctcagcctc 119
|||||
OY 1119 accatgctctcttcgacgaccttgacctcttgaggagctcaagtgatcctcctgcctcagcttc 1178
|||||

Db 120 ctgagtagctgagcactccacacacacacagcagcctgctgactaattttgatttttt 179
|||||
OY 1179 CT-AGTACCTGGAAGT-ACAAGGAGGGGCGACACACCGTAGCTAATCTTTTGTGTTT 1236
|||||

Db 180 ttgt-tagagatggggtttcaccatgtttgcccagctgtctc 221
|||||
OY 1237 TTTGGTAAAGATGGCATTTCGCCGATGTGTACAGGCTGGTCTC 1279
|||||

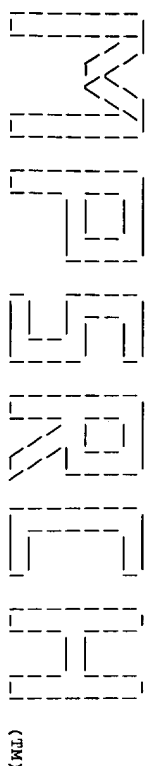
RESULT 15
LOCUS AA262532 300 bp mRNA EST 12-JUN-1997
DEFINITION zsl7c02.t1 NCI-CGAP_GCH1 Homo sapiens CDNA clone 685442 5' similar
to contains Alu repetitive element;contains element MSRI repetitive
element;.
ACCESSION AA262532
NID g1897894
KEYWORDS EST.
SOURCE human

```

ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 300)
AUTHORS	NCI-CGAP.
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert_Strausberg@nih.gov This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Putative full length read The vector to vector length is Seq primer: -28ml3 rev2 E7 from Amersham.
FEATURES	Location/Qualifiers 1..300
SOURCE	/organism="Homo sapiens" /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+ IgD-), generated by Dr. Louis M. Staudt (NCL), Dr. David Allman (NCL) and Dr. Gerald Marti (CEBR). cDNA synthesis was primed with a Not I - 0190(dt) primer 15'-GTGTACCAATCTGAGTGGAGAGCGCCGCTCATTTTTTTTTTTTTT-3'). Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo." /clone_1id="685442" /clone_1id="NCI_CGAP_GCB1" /tissue_type="germinal center B cell" /lab_host="DH10B" <1..>300
BASE COUNT	86 a 86 t 77 g 77 c
ORIGIN	mrna
Query Match	7.6%; Score 108; DB 20; Length 300;
Best Local Similarity	76.2%; Pred. No. 1,56e-59;
Matches	179; Conservative 0; Mismatches 53; Indels 3; Gaps 3;
Db	68 tggagagactcttgagagccagaattttagagaccaccttgggcaagatagtgaacccta 127
Cp	1307 TGGGAGGCCAAGAGGAACCTTAGAGATTGAGAACACAGCCTGTACAAACATGGGGAATGCA 1248
Db	128 tctctataaa-aaaaataaaaaataatgacccaggcgtgaatgagggcacacctgtaattcc 186
Cp	1247 TCTTATCCAAACAAACAAAAAAGATAGTCAGAGTGCTGGTGGCCCTTCTTGAGTTCC 1188
Db	187 agctacacagaagagctgaggaagaggttacttgaagc-cgaagagatcgaagctgtatc 245
Cp	1187 AGCTACT-AGAGAGCTGAGGCGAGGAGATCTCTTGAGCTCCCAAGGTCAAGAGCTGACA 1129
Db	246 gagctatgatgagggccactgcacctaagcctctggcgatagatagatagacacctgtca 300
Cp	1128 GAGCGATGTGTGTGCACATGCACATCCACACCTGGGGAGACAGTGTGAGACCCGTCA 1074

Search completed: Wed Mar 25 14:03:35 1998
 500 time : 532 secs.

THIS PAGE BLANK (USPTO)



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MSPrch_n n.a. - n.a. database search, using Smith-Waterman algorithm
MasPar time 76.15 Seconds
854.742 Million cell updates/sec

Title: >US-08-236-918A-7
(1-1415) from 5674704.seq
Description: 1415
Perfect Score: 1415
N.A. Sequence: 1 AGTGGAAAGTCTCCGCGAG.....TGATTAATAAAAAAAAAA 1415
Comp: TCACCTTCAAGAGCCGTC.....ACTATTTTTTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 87538 seqs, 22999015 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

n-issued
1:back1 2:51 3:52 4:53 5:54 6:55 7:56 8:57 9:PCPT90
10:PCPT91 11:PCPT92 12:PCPT93 13:PCPT94 14:PCPT95 15:PCPT96

Statistics: Mean 9.033; Variance 5.543; scale 1.630

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1415	100.0	1415	7	US-08-236- Sequence 7, Applicatio	0.00e+00
2	836	59.1	838	13	PCPT-US94-1 Sequence 1, Applicati	0.00e+00
3	836	59.1	838	15	PCPT-US96-0 Sequence 7, Applicatio	0.00e+00
4	233	16.5	2347	15	PCPT-US96-0 Sequence 1, Applicatio	1.35e-135
5	229	16.2	768	7	US-08-236- Sequence 5, Applicatio	7.20e-133
6	105	7.4	6769	14	PCPT-US95-1 Sequence 20, Applicati	9.80e-50
7	105	7.4	6769	14	PCPT-US95-1 Sequence 20, Applicati	9.80e-50
8	105	7.4	6769	14	PCPT-US95-1 Sequence 20, Applicati	9.80e-50
9	105	7.4	6769	14	PCPT-US95-1 Sequence 20, Applicati	9.80e-50
10	103	7.3	2048	5	US-08-261- Sequence 11, Applicati	1.96e-48
11	103	7.3	2048	4	US-07-602- Sequence 11, Applicati	1.96e-48
12	101	7.1	7210	14	PCPT-US95-0 Sequence 10, Applicati	3.91e-47
13	101	7.1	22481	14	PCPT-US95-0 Sequence 43, Applicati	3.91e-47
14	99	7.0	20303	7	US-08-370- Sequence 6, Applicatio	7.75e-46
15	99	7.0	26764	7	US-08-370- Sequence 1, Applicatio	1.53e-46
16	97	6.9	2362	6	US-08-265- Sequence 3, Applicatio	3.44e-45
17	98	6.9	5095	5	US-08-092- Sequence 3, Applicatio	3.44e-45
18	98	6.9	19011	7	US-08-310- Sequence 36, Applicati	3.44e-45
19	98	6.9	19557	11	PCPT-US92-0 Sequence 1, Applicatio	3.44e-45

c	20	96	6.8	3090	13	PCT-US94-0	Sequence 3, Applicatio	6.77e-44
c	21	96	6.7	31571	7	US-08-323-	Sequence 1, Applicatio	6.77e-44
c	22	95	6.7	7210	14	PCT-US95-0	Sequence 10, Applicati	2.99e-43
c	23	95	6.7	22481	14	PCT-US95-0	Sequence 43, Applicati	2.99e-43
c	24	94	6.6	1323	15	PCT-US96-1	Sequence 15, Applicati	1.32e-42
c	25	93	6.6	1804	12	PCT-US93-0	Sequence 82, Applicati	5.83e-42
c	26	94	6.6	2619	15	PCT-US96-1	Sequence 17, Applicati	1.32e-42
c	27	94	6.6	2887	15	PCT-US96-1	Sequence 14, Applicati	1.32e-42
c	28	94	6.6	4421	14	PCT-US95-0	Sequence 9, Applicatio	2.57e-41
c	29	92	6.5	1545	7	US-08-446-	Sequence 6, Applicatio	2.57e-41
c	30	92	6.5	1545	7	US-08-446-	Sequence 4, Applicatio	2.57e-41
c	31	92	6.5	5836	7	US-08-380-	Sequence 1, Applicatio	2.57e-41
c	32	92	6.5	5836	7	US-08-380-	Sequence 15, Applicati	2.57e-41
c	33	92	6.5	17327	4	US-07-906-	Sequence 19, Applicati	4.96e-40
c	34	90	6.4	1301	15	PCT-US96-1	Sequence 21, Applicati	4.96e-40
c	35	90	6.4	1363	14	PCT-US95-0	Sequence 33, Applicati	4.96e-40
c	36	90	6.4	1443	15	PCT-US96-1	Sequence 1, Applicatio	1.13e-40
c	37	91	6.4	3590	7	US-08-587-	Sequence 1, Applicatio	1.13e-40
c	38	91	6.4	3590	15	PCT-US96-0	Sequence 20, Applicati	4.96e-40
c	39	90	6.4	6769	14	PCT-US95-1	Sequence 20, Applicati	4.96e-40
c	40	90	6.4	6769	14	PCT-US95-1	Sequence 20, Applicati	4.96e-40
c	41	90	6.4	6769	14	PCT-US95-1	Sequence 20, Applicati	4.96e-40
c	42	90	6.4	6769	7	US-08-480-	Sequence 20, Applicati	4.96e-40
c	43	90	6.4	15328	13	PCT-US94-0	Sequence 33, Applicati	4.96e-40
c	44	91	6.4	17327	4	US-07-906-	Sequence 15, Applicati	1.13e-40
c	45	91	6.4	35100	12	PCT-US93-0	Sequence 19, Applicati	1.13e-40

ALIGNMENTS

RESULT 1
ID US-08-236-918A-7 STANDARD; DNA; UNC; 1415 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 7, Application US/08236918A.
CC Sequence 7, Application US/08236918A
CC Patent No. 5674704
CC GENERAL INFORMATION:
CC APPLICANT: Alderson, Mark R.
CC APPLICANT: Goodwin, Raymond G.
CC APPLICANT: Smith, Craig A.
CC TITLE OF INVENTION: Cytokine Designated 4-1BB Ligand
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
CC STREET: 51 University Street
CC CITY: Seattle
CC STATE: Washington
CC COUNTRY: US
CC ZIP: 98101
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: Apple Power Macintosh
CC OPERATING SYSTEM: Apple 7.5.3
CC SOFTWARE: Microsoft Word, Version #6.0.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/236,918A
CC FILING DATE: 06-May-1994
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/060,843
CC FILING DATE: 07-May-1993
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Anderson, Kathryn A.
CC REGISTRATION NUMBER: 32,172
CC REFERENCE/DOCKET NUMBER: 2801-B
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (206) 587-0430
CC TELEFAX: (206) 233-0644
CC INFORMATION FOR SEQ ID NO: 7:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1415 base pairs

CC TYPE: nucleic acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: cDNA to mRNA
 CC HYPOTHEICAL: NO
 CC ANTI-SENSE: NO
 CC IMMEDIATE SOURCE:
 CC CLONE: hu4-1BB
 CC FEATURE:
 CC NAME/KEY: CDS
 CC LOCATION: 120..887
 CC FEATURE:
 CC NAME/KEY: mat.peptide
 CC LOCATION: 189..884
 CC FEATURE:
 CC NAME/KEY: sig.peptide
 CC LOCATION: 120..188
 CC
 SQ Sequence 1415 BP; 385 A; 332 C; 333 G; 365 T; 0 other;

Query Match 100.0%; Score 1415; DB 7; Length 1415;

Local Similarity 100.0%; Pred.No. 0.00e+00; Mismatches 0; Indels 0; Gaps 0;

1 AGTGAAGAGTCTCCGCGAGCCCTGAGATCTCAAGAGTGAATTTGTGAGACCAAGTAA 60
 1 AGTGAAGAGTCTCCGCGAGCCCTGAGATCTCAAGAGTGAATTTGTGAGACCAAGTAA 60
 61 TTGATTAATTTCTTCTTGAATCAGCTTTGCTAGTATACCTGTGCGAATTTCA 120
 61 TTGATTAATTTCTTCTTGAATCAGCTTTGCTAGTATACCTGTGCGAATTTCA 120
 121 TGGGAACAGCTGTACAAATAGTACGACCTGTGCTGCTCCTCAACTTTGAGAGA 180
 121 TGGGAACAGCTGTACAAATAGTACGACCTGTGCTGCTCCTCAACTTTGAGAGA 180
 121 TGGGAACAGCTGTACAAATAGTACGACCTGTGCTGCTCCTCAACTTTGAGAGA 180
 121 TGGGAACAGCTGTACAAATAGTACGACCTGTGCTGCTCCTCAACTTTGAGAGA 180
 181 CAAGATCATTCAGAGATCTCTGTAGTAACTGCCAGCTGTACATTTCTGTATATA 240
 181 CAAGATCATTCAGAGATCTCTGTAGTAACTGCCAGCTGTACATTTCTGTATATA 240
 181 CAAGATCATTCAGAGATCTCTGTAGTAACTGCCAGCTGTACATTTCTGTATATA 240
 241 GGAATCAGATTTGCAAGTCCCTGCTCCTCAAAATAGTTTCCAGCGCAGAGTGA 300
 241 GGAATCAGATTTGCAAGTCCCTGCTCCTCAAAATAGTTTCCAGCGCAGAGTGA 300
 241 GGAATCAGATTTGCAAGTCCCTGCTCCTCAAAATAGTTTCCAGCGCAGAGTGA 300
 301 CCTGTGACATATGAGAGGAGTGAAGGTGTTTTCAGGACGAGAGAGAGTGTCT 360
 301 CCTGTGACATATGAGAGGAGTGAAGGTGTTTTCAGGACGAGAGAGAGTGTCT 360
 301 CCTGTGACATATGAGAGGAGTGAAGGTGTTTTCAGGACGAGAGAGAGTGTCT 360
 361 CCAGCAATGACAGAGTGTGATGCACTCCAGAGTTTCACTGCTGAGGAGCAGAGT 420
 361 CCAGCAATGACAGAGTGTGATGCACTCCAGAGTTTCACTGCTGAGGAGCAGAGT 420
 361 CCAGCAATGACAGAGTGTGATGCACTCCAGAGTTTCACTGCTGAGGAGCAGAGT 420
 421 TGTGTGAACAGAGTGTGATGCACTCCAGAGTTTCACTGCTGAGGAGCAGAGT 480
 421 TGTGTGAACAGAGTGTGATGCACTCCAGAGTTTCACTGCTGAGGAGCAGAGT 480
 421 TGTGTGAACAGAGTGTGATGCACTCCAGAGTTTCACTGCTGAGGAGCAGAGT 480
 481 GCTTGGGACATTTAAAGCATGAGAAAGTGTGATGCACTCCAGAGTTTCACT 540
 481 GCTTGGGACATTTAAAGCATGAGAAAGTGTGATGCACTCCAGAGTTTCACT 540
 481 GCTTGGGACATTTAAAGCATGAGAAAGTGTGATGCACTCCAGAGTTTCACT 540
 481 GCTTGGGACATTTAAAGCATGAGAAAGTGTGATGCACTCCAGAGTTTCACT 540
 541 TGTGTGAACAGAGTGTGATGCACTCCAGAGTTTCACTGCTGAGGAGCAGAGT 600
 541 TGTGTGAACAGAGTGTGATGCACTCCAGAGTTTCACTGCTGAGGAGCAGAGT 600
 541 TGTGTGAACAGAGTGTGATGCACTCCAGAGTTTCACTGCTGAGGAGCAGAGT 600
 541 TGTGTGAACAGAGTGTGATGCACTCCAGAGTTTCACTGCTGAGGAGCAGAGT 600
 601 CTCGAGCGAGCTCTCTCCGAGAGCATCTGTGACCCCGCTGCGCTCGAGAGAGC 660
 601 CTCGAGCGAGCTCTCTCCGAGAGCATCTGTGACCCCGCTGCGCTCGAGAGAGC 660
 601 CTCGAGCGAGCTCTCTCCGAGAGCATCTGTGACCCCGCTGCGCTCGAGAGAGC 660
 601 CTCGAGCGAGCTCTCTCCGAGAGCATCTGTGACCCCGCTGCGCTCGAGAGAGC 660
 661 CAGAGACATCTCCGAGAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
 661 CAGAGACATCTCCGAGAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
 661 CAGAGACATCTCCGAGAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
 721 TCTGTGATCT 780
 721 TCTGTGATCT 780

QY 721 TCTGTGATCT 780
 DB 781 TGTATATATCAACCAACCACTTTTGTGAGACGATACCACTCTCAAGAGAAATGGCT 840
 QY 781 TGTATATATCAACCAACCACTTTTGTGAGACGATACCACTCTCAAGAGAAATGGCT 840
 DB 841 GTAGCTCCGATTTCCAGAGAAAGAGAGAGATGTGAATCTGAAATGGAATCAAT 900
 QY 841 GTAGCTCCGATTTCCAGAGAAAGAGAGAGATGTGAATCTGAAATGGAATCAAT 900
 DB 901 AGGCGTGTGGGACTTTCTTGAAGAAAGAGAGAAATATGATCATCCGCTATCACAGC 960
 QY 901 AGGCGTGTGGGACTTTCTTGAAGAAAGAGAGAAATATGATCATCCGCTATCACAGC 960
 DB 961 TTTCAAAAGCAAGAACCACTCTCAATATACCAAGATTTCCCCCAACACAGCTTTT 1020
 QY 961 TTTCAAAAGCAAGAACCACTCTCAATATACCAAGATTTCCCCCAACACAGCTTTT 1020
 DB 1021 TCTAATGCCAATGAGTGGCTTTAAATGACCACTTTTCTTTTCTTTTGTGACAGG 1080
 QY 1021 TCTAATGCCAATGAGTGGCTTTAAATGACCACTTTTCTTTTCTTTTGTGACAGG 1080
 DB 1081 GTCTCACTCTGTACCCAGGCTGAGTGCAGTGGCACACCATGCTCTGACAGCCTTG 1140
 QY 1081 GTCTCACTCTGTACCCAGGCTGAGTGCAGTGGCACACCATGCTCTGACAGCCTTG 1140
 DB 1141 ACCTCTGGAGCTCAATGATCTCTCTGCTCCTCACTCTCTCTCTCTCTCTCTCTCT 1200
 QY 1141 ACCTCTGGAGCTCAATGATCTCTCTGCTCCTCACTCTCTCTCTCTCTCTCTCTCT 1200
 DB 1201 AGGCGCACACACCTGACTAATTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTT 1260
 QY 1201 AGGCGCACACACCTGACTAATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1260
 DB 1261 TGTGTGACAGGCTGTCTCAAACTCTGAGTCTCACTTTTGGCTCCCAAGTGTGGATT 1320
 QY 1261 TGTGTGACAGGCTGTCTCAAACTCTGAGTCTCACTTTTGGCTCCCAAGTGTGGATT 1320
 DB 1321 ACAGACATGAACTGCCAGGCGCCGCAAAATTAATGACCACTTTTAAACAGAGACAGA 1380
 QY 1321 ACAGACATGAACTGCCAGGCGCCGCAAAATTAATGACCACTTTTAAACAGAGACAGA 1380
 DB 1381 TGAGGACAGAGCTGTGATATAAAAAAAAAAAAAA 1415
 QY 1381 TGAGGACAGAGCTGTGATATAAAAAAAAAAAAAA 1415

RESULT 2
 ID PCT-US94-10457-1 STANDARD; DNA; UNC; 838 BP.
 AC xxxxxx
 DT 01-JAN-1900
 DE Sequence 1, Application PC/TUS9410457.
 CC Sequence 1, Application PC/TUS9410457
 CC GENERAL INFORMATION:
 CC APPLICANT: Byoung Se Kwon
 CC TITLE OF INVENTION: New Human Receptor and Related Products
 CC NUMBER OF SEQUENCES: 1
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Barnard & Brown
 CC STREET: 306 E. State St., Suite 220
 CC CITY: Ithaca
 CC STATE: New York
 CC COUNTRY: United States
 CC ZIP: 14850
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
 CC COMPUTER: IBM AT Compatible
 CC OPERATING SYSTEM: MS DOS, Version 5.0
 CC SOFTWARE: Special Qbasic program
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: PCT/US94/10457
 CC FILING DATE:


```

CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/03965
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/122,796
CC FILING DATE: 16-SEP-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/012,269
CC FILING DATE: 01-FEB-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 996
CC FILING DATE: 30-JUL-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/267,577
CC FILING DATE: 07-NOV-1988
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Michaels, Christopher A
CC REGISTRATION NUMBER: 34,390
CC REFERENCE/DOCKET NUMBER: KMO5
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 607-273-1711
CC TELEFAX: 607-273-2609
CC INFORMATION FOR SEQ ID NO: 7:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 838 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to mRNA
CC HYPOTHEICAL: NO
CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC ORGANISM: Homo sapiens
CC INDIVIDUAL ISOLATE: H4-1BB #1
CC DEVELOPMENTAL STAGE: Differentiated T-cell
CC CELL TYPE: Lymphocyte
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 41..805
CC OTHER INFORMATION: /codon_start= 41
CC OTHER INFORMATION: /product= "H4-1BB"
CC OTHER INFORMATION: /number= 1
CC FEATURE:
CC NAME/KEY: mat_peptide
CC LOCATION: 41..802
CC OTHER INFORMATION: /codon_start= 41
CC OTHER INFORMATION: /product= "H4-1BB"
CC OTHER INFORMATION: /number= 1
CC Sequence 838 BP; 218 A; 191 C; 215 G; 214 T; 0 other;
CC
CC Query Match 59.1%; Score 836; DB 15; Length 838;
CC Best Local Similarity 99.9%; Pred. No. 0.00e+00;
CC Matches 837; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CC
CC 1 AATCAGCTTGTCTAGTATCATACCTCTGTCGCCAGATTTCATCATGGGAAACAGCTGTACAA 60
CC |||||||
CC 80 AATCAGCTTGTGTAGTATCATACCTCTGTCGCCAGATTTCATCATGGGAAACAGCTGTACAA 139
CC |||||||
CC 61 CATAGTACCACTCTGTTGCTGTCCTTAACCTTTGAGAGCAAGATCATTCAGATCC 120
CC |||||||
CC 140 CATAGTACCACTCTGTTGCTGTCCTTAACCTTTGAGAGCAAGATCATTCAGATCC 199
CC |||||||
CC 121 TTGAGTAGACGCCAGCGTGTACTCTTGATTAATTAACGGAATCAGATTTCAGATCC 180
CC |||||||
CC 200 TTGAGTAGACGCCAGCGTGTACTCTTGATTAATTAACGGAATCAGATTTCAGATCC 259
CC |||||||
CC 181 CTGTCTTCAATATAGTTTCTCCAGCGCAGGTGGACAAGACCTGTGCATATCAGAGCA 240
CC |||||||
CC 260 CTGTCTTCAATATAGTTTCTCCAGCGCAGGTGGACAAGACCTGTGCATATCAGAGCA 319
CC |||||||
CC 241 GTGTAAAGGTGTTTTCAGGACCGAAGAGAGTGTCTCCACCAAGCATGCAAGATGCA 300
CC |||||||

```

QY	320	GTGTAAAGGTGTTTTTCAGSAGCAGGAAAGAGAGTGTCTCCACCAAGCAATGCAGAGTGTGA	379
Db	301	CTGCACCTCCAGGGTTTCACTGCTCTGGGGGCGAGATGCAGCATGTGTGAACAGGATTGTAA	360
QY	380	CTGCACCTCAGGGTTTCACTGCTCTGGGGGCGAGATGCAGCATGTGTGAACAGGATTGTAA	439
Db	361	ACAAGCTCAACACATCGACAAAAAAAGTTGTAAGACTGTGTGGACATTAAACA	420
QY	440	ACAAAGTCAAAACATCGACAAAAAAAAGTTGTAAGACTGTGTGGACATTAAACA	499
Db	421	TCAGAAAGCTGGCATCTGTCCAGCCCTGGACAAATGTGTCTTTGGATGAAAGTGTGCT	480
QY	500	TCAGAAAGCTGGCATCTGTCCAGCCCTGGACAAATGTGTCTTTGGATGAAAGTGTGCT	559
Db	481	TGTGAATGGGACGAAAGAGAGAGGAGCTGTGTCTGTGGAACATCTCCAGCTGAACCTCTCC	540
QY	560	TGTGAATGGGACGAAAGAGAGAGGAGCTGTGTCTGTGGAACATCTCCAGCCACCTCTCC	619
Db	541	GGGACATCTCTGTGTGACATCCCGCTGCCCCCTGCGACAGAGAGCCAGGAACATCTCCGAGAT	600
QY	620	GGGACATCTCTGTGTGACATCCCGCTGCCCCCTGCGACAGAGAGCCAGGAACATCTCCGAGAT	679
Db	601	CATCCTCTTCTTCTGTGGGCTGAGAGTGCAGCTGCTCTCTCTGCTGTCTTCCACAC	660
QY	680	CATCCTCTTCTTCTGTGGGCTGAGAGTGCAGCTGCTGCTCTCTCTGCTGTCTTCCACAC	738
Db	661	GCTCCGTTTCTCTGTGTGTAACGCGGCGAGAAAGAACTCTGTATATATTCAAACACAC	720
QY	740	GCTCCGTTTCTCTGTGTGTAACGCGGCGAGAAAGAACTCTGTATATATTCAAACACAC	799
Db	721	ATTATATGACACCACTACAACTACTCAAGAGGAAGATGCTGTAGCTGCCGATTTCACA	780
QY	800	ATTATATGACACCACTACAACTACTCAAGAGGAAGATGCTGTAGCTGCCGATTTCACA	859
Db	781	AGAAGAAAGAGAGATGTGAACCTGTGAATGGAAGTCAATAGAGGCTGTGGGACTTT	838
QY	860	AGAAGAAAGAGAGATGTGAACCTGTGAATGGAAGTCAATAGAGGCTGTGGGACTTT	917
RESULT 4			
ID	PCT-US96-03965-1	STANDARD; DNA; UNC; 2347 BP.	
AC	xxxxxx		
DT	01-JAN-1900		
DE	Sequence 1, Application PC/TUS9603965.		
CC	Sequence 1, Application PC/TUS9603965		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Kwon, Byoung Se		
CC	TITLE OF INVENTION: Monoclonal antibody against human		
CC	TITLE OF INVENTION: receptor 4-1BB		
CC	NUMBER OF SEQUENCES: 10		
CC	CORRESPONDENCE ADDRESS:		
CC	ADDRESSEE: Barnard, Brown & Michels		
CC	STREET: 306 East State Street, Suite 220		
CC	CITY: Ithaca		
CC	STATE: NY		
CC	COUNTRY: USA		
CC	ZIP: 14850		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: IBM PC compatible		
CC	OPERATING SYSTEM: PC-DOS/MS-DOS		
CC	SOFTWARE: PatentIn Release #1.0, Version #1.30		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: PCT/US96/03965		
CC	FILING DATE:		
CC	CLASSIFICATION:		
CC	PRIOR APPLICATION DATA:		
CC	APPLICATION NUMBER: US 08/122,796		
CC	FILING DATE: 16-SEP-1993		
CC	PRIOR APPLICATION DATA:		
CC	APPLICATION NUMBER: US 08/012,269		
CC	FILING DATE: 01-FEB-1993		


```
CC IMMEDIATE SOURCE:
CC CLONE: mu4-1B8
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..768
CC FEATURE:
CC NAME/KEY: mat_peptide
CC LOCATION: 70..768
CC FEATURE:
CC NAME/KEY: sig_peptide
CC LOCATION: 1..69
CC
CC Sequence 768 BP: 188 A, 186 C, 217 G, 177 T, 0 other
```

Query Match	16.28;	Score 229;	DB 7;	Length 768;
Best Local Similarity	69.08;	Pred. No. 7.20e-133;		
Matches 524;	Conservative 0;	Mismatches 223;	Indels 12;	Gaps 8

D	b	1	ATGGGAAACACCTGTTTAAACAGGTGGTGGCTCATTTGGCTGTGCTAGTGGGCTGGCAGAG	60
Q	y	120	ATGGGAAACACCTGTTTAAACATAGTACCACTGTGTTGGTGGCTCAACTTTGAGAGG	179
		61	GTGGAGCCCGTCGACAGACTCCTGTGATTAACGTCTACGCTGGTACTTTTGC-AGAA-AAT	118
Q	y	180	ACAAGATATTGTCAGGAGCTCTTGTAAGTACGCCAGCTGGTACATTTGTGATAAATAAC	239
D	b	119	ACAATCCAG-TCTGCAGAGCTGCCCTCCAGTACCTCTCCAGACTAGTGGAGACGCC	177
Q	y	240	AGGAATCAAGATTTGGAGTCCCTGTCTCTCCAAATATGTTTCTCCAGCGCAGGTGGACAAAG	299
D	b	178	AACTGTAACTCTGCAGAGTGTGTGCAGGCTAATTTACAGTTTCAAGAGTTTGTCTCTCT	237
Q	y	300	ACCTGTGCAATATCCAGGCACTGTAAAGGTGTTTTCAGAGACAGAAAGAGTGTCTCC	359
D	b	238	ACCCACAACGGGAGTGTGATGCAATTAAGGATTCATCTCTTGGGGCCCAAGTGCACC	297
Q	y	360	ACCAACCAATTCAGAGATGTGACTGCATCTCCAGGTTTACCTCTGGGGGCGAGATGCAGC	419
D	b	298	AGATGTGAAAAGACCTGCAGGCTTGCCAGAGCTAACGAAGCAGGGTTGCAAAACCTGT	357
Q	y	420	ATGTGTGACACAGGATTTGTAAACAAGGTCCAAACACTGCACAAAAAAGGTGTAAAGACTGT	479
D	b	358	AGCTTGGGACATTTAATGACACAGAGGTACTGGCGTGTGCACCTTGGACGACACTGC	417
Q	y	480	TGCTTTGGGACATTTAATGACATCAGAAAGCT---GGCATCTGTGACCTTGGACAAACTGT	536
D	b	418	TCTCTAGACGGAAGTCTGTGCTTAAGACCGGACCAACGAGAAAGAGTGGTGTGTGA	477
Q	y	537	TCTTTGATGGAAGTCTGTGCTTGTGAATGAGACGAAGAGAGAGAGTGTGTGTGA	596
		478	CCCCCTGTGCTCTCTCTCC--AGTA-CCACCAATTTGTGTGATCTCCAGAGGAGAGA	534
Q	y	597	CCATCTCCAGCGCACTCTCTCCGGGACCACTCTGTGACCCCGCTCCCGCTCCGAGA	656
D	b	535	CCAGAGAGGCACTCTTCGACAGTCTCTTCACTGTTCCTGGCGCTGACATCGGCTTTCGTG	594
Q	y	657	GAGCGAGGACACTCTCCGACATCATCTCTCTTCTTGTGGCGTGAAGTGCAGTGCAGT	716
D	b	595	CTGG-CC--CTGATCTTCACTACTCTCTGTCTCTGTGTCTGTCTCAATGATAGCAAAAA	651
Q	y	717	CTCTTCGCTGCTCTCTCTCCACCTCCGCTTCTCTGTGTTAAACGGGCGAGAAAGAAA	776
D	b	652	TTCCCCCACTATTCAAGCAACATTTAAGAAGCACTGGAGCAGCTCAAGAGAGAGAT	711
Q	y	777	CTCTGTATATATTCACAAACACATTTATGAGACCAGTACAAACTACTCAAGAGAGAGAT	836
D	b	712	GCTTGTAGCTGCGATGTCCACAGAAAGAGAGAGAGA	750
Q	y	837	GGCTGTAGCTGCGATTTCCAGAAAGAGAGAGAGAGA	875

```

DE 01-JAN-1900
CC Sequence 20, Application PC/TUS9510302.
CC Sequence 20, Application PC/TUS9510202.
CC GENERAL INFORMATION:
CC APPLICANT: Shattuck-Eidens, Donna M.
CC APPLICANT: Simard, Jacques
CC APPLICANT: Emi, Mitsuru
CC APPLICANT: Nakamura, Yusuke
CC APPLICANT: Durocher, Francine
CC TITLE OF INVENTION: In vivo Mutations and Polymorphisms
CC TITLE OF INVENTION: In the 17q-Linked Breast and Ovarian Cancer
CC NUMBER OF SEQUENCES: 85
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
CC STREET: 1201 New York Avenue, N.W., Suite 1000
CC CITY: Washington
CC STATE: DC
CC COUNTRY: USA
CC ZIP: 20005
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/10202
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US
CC FILING DATE: 07-JUN-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/409,305
CC FILING DATE: 24-MAR-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/348,824
CC FILING DATE: 29-NOV-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08-308,104
CC FILING DATE: 16-SEP-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/300,266
CC FILING DATE: 02-SEP-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/289,221
CC FILING DATE: 12-AUG-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Ihnen, Jeffrey L.
CC REGISTRATION NUMBER: 28,957
CC REFERENCE/DOCKET NUMBER: 24884-109347
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 202-962-4810
CC TELEFAX: 202-962-8300
CC INFORMATION FOR SEQ ID NO: 20:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 6769 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC ORGANISM: Homo sapiens
CC Sequence 6765 BP: 1954 A; 1377 C; 1469 G; 1937 T; 32 other;
SO
Query Match 7.4%; Score 105; DB 14; Length 6769;
Best Local Similarity 76.1%; Pred. No. 9.80e-50;
Matches 175; Conservative 0; Mismatches 52; Indels 3; Gaps 3

```


Cp 1286 GAGATTGAGACGACCTGTACACATGCGAAATGCCATCTTACCAACAAAAACAA 1227
 Db 2850 AAAAAATTACCCGGGGTGGTGGCTATGCTGTATCCAGCTACTAGAGAGGCTGAGGC 2909
 Cp 1226 AAAAGATTAGTCAGGTGGTGGCCCTCTTGTAGTTCAGCTACT-AGGAGACTGAGGC 1168
 Db 2910 AGGAGATCCCTTGAAC-CCAGAGAGCAGAGGTGTGACGTGAGCAACATGCGACCATTTGC 2968
 Cp 1167 AGGAGATCCTTGTAGCTCCAGAGGTCAAGGCTGCGAGAGCCATGTGTGCTGCTGC 1108
 Db 2969 ACTCCAGCCTAGGACACAGAGTGAATCCATCTCAAAAAAAAAAAAAA 3018
 Cp 1107 ACTCCAGCCTGGTGACA-GAGTGAGACCTGTCAAAAAAAAAAAAAA 1059

RESULT 7
 ID PCT-US95-10203-20 STANDARD; DNA; UNC; 6769 BP.
 AC xxxxxx
 DT 01-JAN-1900
 Sequence 20, Application PC/TUS9510203.
 Sequence 20, Application PC/TUS9510203.
 GENERAL INFORMATION:
 CC APPLICANT: Skolnick, Mark H.
 CC APPLICANT: Goldgar, David E.
 CC APPLICANT: Miki, Yoshio
 CC APPLICANT: Swenson, Jeff
 CC APPLICANT: Kamb, Alexander
 CC APPLICANT: Harshman, Keith D.
 CC APPLICANT: Shattuck-Eidens, Donna M.
 CC APPLICANT: Tavtigian, Sean V.
 CC APPLICANT: Wiseman, Roger W.
 CC APPLICANT: Futreal, P. Andrew
 CC TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
 CC TITLE OF INVENTION: Susceptibility Gene
 CC NUMBER OF SEQUENCES: 85
 CC CORRESPONDENCE ADDRESSES:
 CC ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
 CC STREET: 1201 New York Avenue, N.W., Suite 1000
 CC CITY: Washington
 CC STATE: DC
 CC COUNTRY: USA
 CC ZIP: 20005
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: PCT/US95/10203
 CC FILING DATE:
 CC CLASSIFICATION:
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US
 CC FILING DATE: 07-JUN-1995
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/409,305
 CC FILING DATE: 24-MAR-1995
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/348,824
 CC FILING DATE: 29-NOV-1994
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08-308,104
 CC FILING DATE: 16-SEP-1994
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/300,266
 CC FILING DATE: 02-SEP-1994
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/289,221
 CC FILING DATE: 12-AUG-1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Ihnen, Jeffrey L.
 CC REGISTRATION NUMBER: 28,957
 CC REFERENCE/DOCKET NUMBER: 24884-109347

CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 202-962-4810
 CC TELEFAX: 202-962-8300
 CC INFORMATION FOR SEQ ID NO: 20:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 6769 base pairs
 CC TYPE: nucleic acid
 CC STRANDEDNESS: double
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: DNA (genomic)
 CC HYPOTHEICAL: NO
 CC ANTI-SENSE: NO
 CC ORIGINAL SOURCE:
 CC ORGANISM: Homo sapiens
 CC SEQ Sequence 6769 BP; 1954 A; 1377 C; 1469 G; 1937 T; 32 other;

Query Match 7.4%; Score 105; DB 14; Length 6769;
 Best Local Similarity 76.1%; Pred. No. 9,80e-50;
 Matches 175; Conservative 0; Mismatches 52; Indels 3; Gaps 3;

Db 2790 GAGATTGAGACGACCTGTACACATGAGAGAAACCCATCTCTCAAAAAAAAAA 2849
 Cp 1286 GAGATTGAGACGACCTGTACACATGGGAAATGCACTTTACCAACAAAAACAA 1227
 Db 2850 AAAAAATTACCCGGGGTGGTGGCTTATGCTGTATCCAGCTACTAGAGAGGCTGAGGC 2909
 Cp 1226 AAAAGATTAGTCAGGTGGTGGCCCTCTTGTAGTTCAGCTACT-AGGAGACTGAGGC 1168
 Db 2910 AGGAGATCCCTTGAAC-CCAGAGAGCAGAGGTGTGACGTGAGCAACATGCGACCATTTGC 2968
 Cp 1167 AGGAGATCCTTGTAGCTCCAGAGGTCAAGGCTGCGAGAGCCATGTGTGCTGCTGC 1108
 Db 2969 ACTCCAGCCTAGGACACAGAGTGAATCCATCTCAAAAAAAAAAAAAA 3018
 Cp 1107 ACTCCAGCCTGGTGACA-GAGTGAGACCTGTCAAAAAAAAAAAAAA 1059

RESULT 8
 ID PCT-US95-10220-20 STANDARD; DNA; UNC; 6769 BP.
 AC xxxxxx
 DT 01-JAN-1900
 Sequence 20, Application PC/TUS9510220.
 Sequence 20, Application PC/TUS9510220.
 GENERAL INFORMATION:
 CC APPLICANT: Skolnick, Mark H.
 CC APPLICANT: Goldgar, David E.
 CC APPLICANT: Miki, Yoshio
 CC APPLICANT: Swenson, Jeff
 CC APPLICANT: Kamb, Alexander
 CC APPLICANT: Harshman, Keith D.
 CC APPLICANT: Shattuck-Eidens, Donna M.
 CC APPLICANT: Tavtigian, Sean V.
 CC APPLICANT: Wiseman, Roger W.
 CC APPLICANT: Futreal, P. Andrew
 CC TITLE OF INVENTION: Method for Diagnosing a
 CC TITLE OF INVENTION: Predisposition for Breast and Ovarian Cancer
 CC NUMBER OF SEQUENCES: 85
 CC CORRESPONDENCE ADDRESSES:
 CC ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
 CC STREET: 1201 New York Avenue, N.W., Suite 1000
 CC CITY: Washington
 CC STATE: DC
 CC COUNTRY: USA
 CC ZIP: 20005
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: PCT/US95/10220
 CC FILING DATE:
 CC CLASSIFICATION:

OTHER INFORMATION: sequence for PEDF plus flanking sequences.
 CC Sequence 22481 BP; 5280 A; 5708 C; 6136 G; 5347 T; 10 other;
 SQ

Query Match 7.1%; Score 101; DB 14; Length 22481;
 Best Local Similarity 76.2%; Pred. No. 3, 91e-47;
 Matches 176; Conservative 0; Mismatches 51; Indels 4; Gaps 4;

Db 6067 CTTTCTTTTGTGTAAGAGGGGCTCTCTCTGCGCCCAAGATGAGTGCATGGGTGT 6126
 OY TTTTCTTTTGTGTAAGAGGGGCTCTCTCTGCGCCCAAGATGAGTGCATGGGTGT 6117
 Db 6127 GATCTCAGCTCAGTCAACCTCCGCTCCAG-GTTTAAGTATCTCTGCTCAGACT 6185
 OY TTTTCTTTTGTGTAAGAGGGGCTCTCTCTGCGCCCAAGATGAGTGCATGGGTGT 6117
 Db 1118 CACATGCTCTCTGACCTTGTGAGCTCAAGTATCTCTGCTCAGTCT 1177
 OY TTTTCTTTTGTGTAAGAGGGGCTCTCTCTGCGCCCAAGATGAGTGCATGGGTGT 6117
 Db 6186 CCAAGTACGAGGCTGAGTGCAGGCGCCCAACACACCTGGTAAATTTTGT-TCATT- 6243
 OY TTTTCTTTTGTGTAAGAGGGGCTCTCTCTGCGCCCAAGATGAGTGCATGGGTGT 6117
 Db 1178 CCAAGTACGAGGCTGAGTGCAGGCGCCCAACACACCTGGTAAATTTTGT-TCATT- 6243
 OY TTTTCTTTTGTGTAAGAGGGGCTCTCTCTGCGCCCAAGATGAGTGCATGGGTGT 6117
 Db 6244 TTTAGTAGAGATGGGCTTTCACCGCTGTGACTAGCTGTCTCCAGTCTCT 6294
 OY TTTTCTTTTGTGTAAGAGGGGCTCTCTCTGCGCCCAAGATGAGTGCATGGGTGT 6117
 Db 1237 TTTGTAAGATGGCATTTCCGATGTGTACAGGCTGTCTCAAACTCCT 1287

RESULT 14
 ID US-08-370-975B-6 STANDARD; DNA; UNC; 20303 BP.

AC xxxxxx
 DT 01-JAN-1900
 DE Sequence 6, Application US/08370975B.
 CC Sequence 6, Application US/08370975B
 CC Patent No. 5622851

GENERAL INFORMATION:
 CC APPLICANT: Maley, Frank
 CC APPLICANT: Maley, Gladys F.
 CC APPLICANT: Weiner, Karen X.B.
 CC TITLE OF INVENTION: Human Deoxycytidylate Deaminase Gene
 CC NUMBER OF SEQUENCES: 14
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Nixon, Hargrave, Devans & Doyle
 CC STREET: Clinton Square, P.O. Box 1051
 CC CITY: Rochester
 CC STATE: New York
 CC COUNTRY: USA
 CC ZIP: 14603

COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/370, 975B
 CC FILING DATE: 10-JAN-1995
 CC CLASSIFICATION: 435
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Timian, Susan J.
 CC REGISTRATION NUMBER: 34,103
 CC REFERENCE/DOCKET NUMBER: 20894/80
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (716)263-1636
 CC TELEFAX: (716)263-1600

INFORMATION FOR SEQ ID NO: 6:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 20303 base pairs
 CC TYPE: nucleic acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
 CC POSITION IN GENOME:
 CC CHROMOSOME/SEGMENT: 4q35
 CC Sequence 20303 BP; 5454 A; 4115 C; 5052 G; 5682 T; 0 other;

Query Match 7.0%; Score 99; DB 7; Length 20303;
 Best Local Similarity 80.1%; Pred. No. 7, 75e-46;

Matches 185; Conservative 0; Mismatches 38; Indels 8; Gaps 5;

Db 6226 TTTTCTTTTGTGTAAGAGGGGCTCTCTCTGCGCCCAAGATGAGTGCATGGGTGT 6285
 OY TTTTCTTTTGTGTAAGAGGGGCTCTCTCTGCGCCCAAGATGAGTGCATGGGTGT 6118
 Db 6286 ATCTGGCTCAGTCAACCTCCGCTCCAG-TCAAGCATCTCTCTGCTCAGCTC 6344
 OY TTTTCTTTTGTGTAAGAGGGGCTCTCTCTGCGCCCAAGATGAGTGCATGGGTGT 6118
 Db 6345 GGAGCTGTGTGATTTACAGGAGTGCACACGCGCTGTAA--TTTGTATT- 6401
 OY TTTTCTTTTGTGTAAGAGGGGCTCTCTCTGCGCCCAAGATGAGTGCATGGGTGT 6118
 Db 1179 CTAGTAGTG-CA-ACACAGAGAGGCGCCACACACCTGACTTCTTTTGTGTTG 1236
 OY TTTTCTTTTGTGTAAGAGGGGCTCTCTCTGCGCCCAAGATGAGTGCATGGGTGT 6118
 Db 6402 --TAGTAGAGATGGGCTTTCACCGCTGTGACTAGCTGTCTCCAGTCTCT 6450
 OY TTTTCTTTTGTGTAAGAGGGGCTCTCTCTGCGCCCAAGATGAGTGCATGGGTGT 6118
 Db 1237 TTTGTAAGATGGCATTTCCGATGTGTACAGGCTGTCTCAAACTCCT 1287

RESULT 15
 ID US-08-370-975B-1 STANDARD; DNA; UNC; 26764 BP.

AC xxxxxx
 DT 01-JAN-1900
 DE Sequence 1, Application US/08370975B.
 CC Sequence 1, Application US/08370975B
 CC Patent No. 5622851

GENERAL INFORMATION:
 CC APPLICANT: Maley, Frank
 CC APPLICANT: Maley, Gladys F.
 CC APPLICANT: Weiner, Karen X.B.
 CC TITLE OF INVENTION: Human Deoxycytidylate Deaminase Gene
 CC NUMBER OF SEQUENCES: 14
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Nixon, Hargrave, Devans & Doyle
 CC STREET: Clinton Square, P.O. Box 1051
 CC CITY: Rochester
 CC STATE: New York
 CC COUNTRY: USA
 CC ZIP: 14603

COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/370, 975B
 CC FILING DATE: 10-JAN-1995
 CC CLASSIFICATION: 435
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Timian, Susan J.
 CC REGISTRATION NUMBER: 34,103
 CC REFERENCE/DOCKET NUMBER: 20894/80
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (716)263-1636
 CC TELEFAX: (716)263-1600

INFORMATION FOR SEQ ID NO: 1:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 26764 base pairs
 CC TYPE: nucleic acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
 CC POSITION IN GENOME:
 CC CHROMOSOME/SEGMENT: 4q35
 CC Sequence 26764 BP; 7079 A; 5521 C; 6539 G; 7625 T; 0 other;

Query Match 7.0%; Score 99; DB 7; Length 26764;
 Best Local Similarity 80.1%; Pred. No. 7, 75e-46;
 Matches 185; Conservative 0; Mismatches 38; Indels 8; Gaps 5;

Db 8189 TTTTCTTTTGTGTAAGAGGGGCTCTCTCTGCGCCCAAGATGAGTGCATGGGTGT 8248
 OY TTTTCTTTTGTGTAAGAGGGGCTCTCTCTGCGCCCAAGATGAGTGCATGGGTGT 1118

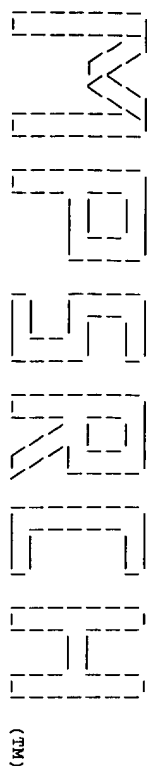
Thu Mar 26 09:16:05 1998

US-08-236-918A-7.rni

Page 12

Db	8249	ATCTTGGCTACTCTCACCCTCCACCTCCAGATT-TCAGAGATTCTCGCTTAGCCTC	8307
Qy	1119	ACCATGGCTCTCTCCACCTTGACCTCTGGAGCTCAAGTGTATCCCTCGCTCAGTCTC	1178
Db	8308	GGAGTAGCTGTGATTACAGAGGACGCGCCACGCGCTGGTAA--TTTGTGATTTT-	8364
Qy	1179	CTAGTAGCTG-GA-ACACAAAGGAGGGCCACACACCTGACACTTTTGTGTTTTTG	1236
Db	8365	--TAGTAGAGATGGGGTTTCCATCATGTTGGCCAGCGCTGCTCAACTCCT	8413
Qy	1237	TTTGTGAAGATGGCATTTTGGCCATGTGTACACGCTGGTCTCAACTCCT	1287

Search completed: Wed Mar 25 14:05:28 1998
Job time : 94 secs.



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Mpsrch_nrn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Wed Mar 25 13:29:04 1998; Maspar time 175.73 Seconds
929.178 Million cell updates/sec
Near output not generated.

Title: >US-08-236-918A-7
Description: (1-1415) from 5674704.seq
Perfect Score: 1415
N.A. Sequence: 1 AGTGGAAAGTCTCCGGCAG.....TCATTAATAAAAAAAAAAAAAA 1415
Comp: TCACCTTCAAGAGCGCGTC.....ACATTTTTTTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 159651 segs, 5769862 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

n-geneseq30
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33

Statistics: Mean 9.526; Variance 6.639; scale 1.435

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	1415	100.0	1415	13	075424	Human 4-1BB polypept	0.00e+00
2	1390	98.2	1439	15	092086	Human receptor induce	0.00e+00
3	836	59.1	838	14	086126	H4-1BB receptor prote	0.00e+00
4	834	58.9	838	24	139546	Human receptor H4-1BB	0.00e+00
5	233	16.5	2347	24	139541	Mouse receptor 4-1BB	7.81e-118
6	231	16.3	2350	14	086127	4-1BB receptor protei	1.24e-115
7	229	16.2	768	13	075428	Murine 4-1BB polypept	1.93e-42
8	114	8.1	2649	6	035034	DNA fragment contg. A	1.34e-47
9	108	7.6	4359	14	094625	Brush-1 cDNA.	3.72e-44
10	108	7.6	11298	33	086756	Human high affinity I	3.72e-44
11	108	7.6	11357	9	051024	Human FCER2 beta chai	3.72e-44
12	105	7.4	24025	21	175151	Mutated BRCA1 genomic	1.93e-42
13	105	7.4	24025	21	175155	Mutated BRCA1 genomic	1.93e-42
14	105	7.4	24026	21	175156	Mutated BRCA1 genomic	1.93e-42
15	105	7.4	24026	21	175159	Mutated BRCA1 genomic	1.93e-42

c	16	105	7.4	24026	21	175152	Mutated BRCA1 genomic	1.93e-42
c	17	105	7.4	24026	21	175157	Mutated BRCA1 genomic	1.93e-42
c	18	105	7.4	24026	21	175151	Mutated BRCA1 genomic	1.93e-42
c	19	105	7.4	24026	21	175156	Mutated BRCA1 genomic	1.93e-42
c	20	105	7.4	24026	21	175150	BRCA1 genomic sequenc	1.93e-42
c	21	105	7.4	24026	21	175154	Mutated BRCA1 genomic	1.93e-42
c	22	105	7.4	24026	21	175153	Mutated BRCA1 genomic	1.93e-42
c	23	105	7.4	24026	21	175158	Mutated BRCA1 genomic	1.93e-42
c	24	105	7.4	24026	21	175151	Mutated BRCA1 genomic	1.93e-42
c	25	105	7.4	24026	21	175152	Mutated BRCA1 genomic	1.93e-42
c	26	105	7.4	24026	21	175153	Mutated BRCA1 genomic	1.93e-42
c	27	105	7.4	24026	21	175159	Mutated BRCA1 genomic	1.93e-42
c	28	105	7.4	24026	21	175154	Mutated BRCA1 genomic	1.93e-42
c	29	105	7.4	24026	18	175157	BRCA1, human breast a	1.93e-42
c	30	105	7.4	24026	21	175157	Mutated BRCA1 genomic	1.93e-42
c	31	105	7.4	24026	23	175152	BRCA1, human breast a	1.93e-42
c	32	105	7.4	24026	21	175150	Mutated BRCA1 genomic	1.93e-42
c	33	105	7.4	24031	21	175155	Mutated BRCA1 genomic	1.93e-42
c	34	104	7.3	230	8	095868	Human brain expressed	7.18e-42
c	35	103	7.3	2045	4	025165	Genomic alpha-GalNAc	2.67e-41
c	36	104	7.3	6511	14	095493	Human Cdn-2 DNA.	7.18e-42
c	37	101	7.1	22481	24	175158	PEPF full length sequ	3.68e-40
c	38	99	7.0	2435	23	175155	Tissue plasminogen ac	5.04e-39
c	39	99	7.0	20303	30	175159	Human deoxyacyldylate	5.04e-39
c	40	99	7.0	26764	30	175156	Human deoxyacyldylate	5.04e-39
c	41	97	6.9	392	8	060438	Human brain expressed	6.88e-38
c	42	97	6.9	2362	26	174809	Human interleukin-12	6.88e-38
c	43	97	6.9	2362	3	023586	Natural killer cell s	1.86e-38
c	44	98	6.9	19012	6	035952	HSA gene.	1.86e-38
c	45	96	6.8	5108	2	160388	Sequence encoding hum	2.34e-37

ALIGNMENTS

RESULT 1
ID 075424; standard: cDNA to mRNA; 1415 BP.
AC 075424;
DE 08-AUG-1995 (first entry)
KW Human 4-1BB polypeptide coding sequence.
KW T-cell; lymphocyte; activation; tissue culture; clone; cell lines;
KW proliferation; stimulation; 4-1BB-L; ligand; receptor; cytokine; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 120..887
FT /product= 4-1BB polypeptide.
FT sig_peptide 120..188
FT /tag= b 189..884
FT mat_peptide
FT /tag= c
FT /tag= c
PN WC09426290-A.
PD 24-NOV-1994.
PF 06-MAY-1994; U05036.
PR 07-MAY-1993; US-060843.
PA (IMV) IMMUNEX CORP.
PI Alderson MR, Goodwin RG, Smith CA;
DR P-PSDB: R64197.
PT Cytokine; 4-1BB ligand (4-1BB-L) - binds to cell surface receptor
PT 4-1BB to transduce signal
PS Claim 29; Page 46-47; 65pp; English.
CC The 4-1BB polypeptide (receptor) and the 4-1BB-L (cytokine, see
CC 075423) are useful in a pharmaceutical composition for stimulating
CC the immune system. The 4-1BB and 4-1BB-L polypeptides are also useful
CC for exploring mechanisms of T-cell activation, as they are expressed
CC on T lymphocytes. 4-1BB-L can be used as a tissue culture reagent for
CC in vitro cultivation of primary T-cells during the derivation of
CC clonal T-cell lines. It may also be used to stimulate proliferation
CC of activated T-cells, used in therapeutic procedures. 365 T;
SQ Sequence 1415 BP; 385 A; 332 C; 333 G; 365 T;
Query Match 100.0%; Score 1415; DB 13; Length 1415;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;

Matches 1415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 1 agtggaaagttctccggcagccctgagatctcaagagtaacattgttagaccagctat 60
QY 1 AGTGGAAAGTTCTCCGGCAGCCCTGAGATCTCAAGATGACATTTGTAGACACCACTAAT 60
Db 61 ttgataaatctctcttgaaatcaagcttgtagatcatatccctgtgcagattcatca 120
QY 61 TTGATTAATAATCTCTTCTTGAAATCAACCTTGGCATAGTATCATACCTTGGCAGATTTCATCA 120
Db 121 tgggaacacagctgttaacaacatagtagcaactgtgtcgtgctcctcaacttgaagga 180
QY 121 TGGGAAACAGCTGTACACATAGTAGCACACTGTCTGCTGCTCCTCAACTTGTAGAGGA 180
Db 181 caagatcttcagagatctctgttagtaactgccagctggtatcatctcgtataataca 240
QY 181 CAAGATCTTCTCAGAGATCTTGTAGTAAGTCCAGCTGTGTATCATCTGTGATAATAACA 240
Db 241 ggaatcaattctgcagctccctgtcccaaatagttctccagcgaggtgacaaagga 300
QY 241 GGAATCAATTTGCAAGTCCCTGTCTCCAAATAGTTTCTCCAGCCGAGTGGACAAAGGA 300
Db 301 ccctgtacatactgcagcagctgttaaggtgtttcaggaaccaaggaaggtgtctcca 360
QY 301 CCTGTGACATATGACAGCAGTGTAAAGTGTTCAGAGACAGAGAGAGTGTCTCTCA 360
Db 361 ccagcaatgcagagatgtgtactgcactccaaggtttcactgctggtggcaggaatgcga 420
QY 361 CCAGCAATGCAGAGATGTACTGCACCTCCAGGATTCACCTGTGCGGCGAGATGCACAGA 420
Db 421 tgtgtgacagagattgtaacaaggtcaagaactgcacaaagagttgttaagagctgt 480
QY 421 TGTGTGAACAGAGATTGTAAACAAGGTCAAGAACTGACAAAAAAGGTTGTAAAGACTTT 480
Db 481 gctttgggacatttaacagatcagaagacgttggacatctgtcgaacctgtgacaaactgt 540
QY 481 GCTTTGGGACATTTAAACATCAGAAACGTGGCATCTGTGCACCCGTGCAACAACCTGTCTT 540
Db 541 tggatggaaatctcgtgtctgttgaatggagaaaggaaggaagcgtgtgtgtgacat 600
QY 541 TGGATGGAATCTCTGTCTGTGTAATGGAGCAAGAGAGGACCTGTGTGTGACCAT 600
Db 601 ctccagccagcactctctccgggagacatcctctgtgaccccgctccctcgagagagac 660
QY 601 CTCACGCCGACCTCTCTCCGGAGACATCTCTGTACCCCGCTCCCTTGGAGAGAGC 660
Db 661 caggacactctccgacagatcatctctcttctctgtcgctgacgtgcagtgtcct 720
QY 661 CAGGACACTCTCCGAGATCATCTCTTCTTCTGCGCTGACGTGCGTGTGCTCT 720
Db 721 tctcgtgtctctcctcagcgtcgttctctctgtgtttaaacggggcaagaatactcc 780
QY 721 TCTCGTGTCTTCTCTCAGCTCCGCTTCTCTGTGTAAACGGGGCAGAAAGAACTCC 780
Db 781 tgtatataccaacaacattatagagaccagtaacaaactcaagaaggaagtgtct 840
QY 781 TGTATATTTTAAACACCAATTTATGAGACCAATGTAACAACTACTAAGAGATGGCT 840
Db 841 gtaactgcagattccagaaagaaagaaagagatgtgaaactgtgaaatgaaatcat 900
QY 841 GTAAGTGCAGATTTCAGAAAGAAAGAAAGAGAGATGTGAATGTGAATGCAAT 900
Db 901 aggcctgttggagcttctctgaaagaagcaagaaataatgagtaaccgctatcaagc 960
QY 901 AGGCTGTGGAGCTTCTCTGAAAAAGAAAGCAAGAAATATGAGTATCCGTATACAGC 960
Db 961 tttaaaaagcaagaaacactactaataataacagagatcccccacaacagttctt 1020
QY 961 TTTAAAAAGCAAGAACACTACTAATAATACCCAGGATTCGCCCAACACAGCTTTT 1020
Db 1021 tctaattgcacatagttgagctttaaaatgcacacatttttttttttttttttgaag 1080
QY 1021 TCTAATGCCAATAGTGTGGCTTTAAAAATGACACACTTTTTTTTTTGTGTGACAGG 1080

```

```

Db 1081 gtctactctgtcacccagcgtgagtgagctgagccacacatgctctctgcagcttg 1140
QY 1081 GTCTACTCTGTCAACCAAGCTGAGATGCAATGACACACCATGCTCTGTGCAAGCTTG 1140
Db 1141 acccttggagactcaatgatactccctgcctcagctcctcagtagctgcgaactacaaga 1200
QY 1141 ACCCTGGAGACTCAATGATCCCTCGGCCATCTCCTAGTAGCTGGAACATCAAGGA 1200
Db 1201 aggcacacacactgactaactttttgtttttgtttgttgaagatgacattgcga 1260
QY 1201 AGGCACACCACTGACTAATCTTTTGTTTTGTGTGTAAGATGGCATTTCCCA 1260
Db 1261 tttgtacagcctgtgtcaaacctcaggtgtacacttggctcccaagtgctggagat 1320
QY 1261 TTTGTACAGCTGTGTCAAACTCTTAGGTTCACTTGGCTCCCAAGTGTGGGATT 1320
Db 1321 acaacacatgaacttccagcggccgccaataatgacacacactttaaacaagacagaga 1380
QY 1321 ACACACATGAACCTCCAGCGCCGCCCAAAATATGACACACTTTAAACAGACAGAGA 1380
Db 1381 tgaggacagagctgtgtataaaaaa 1415
QY 1381 TGAGGACAGAGCTGTGTATAAAAAA 1415

```

RESULT 2

ID 092086 standard; DNA; 1439 BP.

AC 092086;

DT 21-JAN-1996 (first entry)

DE Human receptor induced by lymphocyte activation (ILA) DNA.

KW ILA; receptor inducible by lymphocyte activation; disease diagnosis;

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 140..904

PN /*tag- a

PD 28-MAR-1995. Ca108401-A.

PR 14-OCT-1993; 108401.

PE 27-SEP-1993; US-127693.

PA (REGC) UNIV CALIFORNIA.

PI Lotz M, Schwarz H.

DR WPI: 95-194420/26.

DR P-PSDB: R74087.

PT New receptor inducible by lymphocyte activation - used to develop

PT prods. for the diagnosis and treatment of inflammatory host defence

PT pathology.

PS Claim 52; Page 61; 91pp; English.

CC This DNA may be expressed recombinantly for the production of ILA.

CC The cDNA was isolated from a library constructed from activated

CC human T-lymphocyte leukemia virus type-1 transformed human T-

CC lymphocytes.

SQ Sequence 1439 BP; 393 A; 336 C; 346 G; 364 T;

Query Match 98.2%; Score 1390; DB 15; Length 1439;

Best Local Similarity 99.4%; Pred. No. 0.00e+00;

Matches 1408; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

```

Db 21 agtggaaagttctccggcagccctgagatctcaagagtaacattgttagaccagctat 80
QY 1 AGTGGAAAGTTCTCCGGCAGCCCTGAGATCTCAAGATGACATTTGTAGACCACTAAT 60
Db 81 ttgataaatctctcttgaaatcaagcttgtagatcatatccctgtgcagattcatca 140
QY 61 TTGATTAATAATCTCTTCTTGAAATCAACCTTGGCATAGTATCATACCTTGGCAGATTTCATCA 120
Db 141 tgggaacacagctgttaacaacatagtagcaactgtgtcgtgctcctcaacttgaagga 200
QY 121 TGGGAAACAGCTGTACACATAGTAGCACACTGTCTGCTGCTCCTCAACTTGTAGAGGA 180
Db 201 caagatcttcagagatctctgttagtaactgccagctggtatcatctcgtataataca 260

```


QY 181 CAAGATCATTCAGAGATCTTGTAGTACTGCGCCAGCTGTAACATTTCTGTATATAACA 240
 Db 261 ggaatcagatttgcagtccttgcctcccaaaagtttctccagcaggttgcacaaagga 320
 QY 241 GGAATCAGATTTCAGAGTCCCTGCTCCCAAAATAGTTCTCCAGCCAGGTGACAAAGGA 300
 Db 321 ccttgcacatttgcagcaggttgaaggttcttccagacccaggaaggaagttcttcca 380
 QY 301 CCTGTGACATATGACAGGAGTGTAAAGGTGTTTTCAGACCCAGGAGAGAGTTCTCTCA 360
 Db 381 ccagcaatgcagagatgttgcagtcagtcaggtttcagtgctggggcaggaatgcaga 440
 QY 361 CCAGCAATGCAGAGTGTACTGCTCCAGGTTTCACTGCTGGGGCAGAGATCAGACA 420
 Db 441 tctgtgaacagagatttgcagacaggttgcagacacacaaaggttgaagaagctgt 500
 QY 421 TGTGTGACAGAGATGTAAACAGGTCAAGAACTGACAAAAAAGTTGTAAAGACTGTT 480
 Db 501 gcttgcagacatttgcagacacaggttgcagtcctgcagccctgcagcaactgttctt 560
 QY 481 GCTTGGGACATTTAAACATCAGAAACGTGCATCTGTCGACCCCTGACAAACTGTTCTT 540
 Db 561 tggatggaagaatcgtgtgttgcagtggaaggaaggaaggaagctgtgttgcagcat 620
 QY 541 TGGATGGAAGATCTGTCTGTGGAATGGGACGAGAGAGAGAGACTGGTCTGTGACCAT 600
 Db 621 ctcaagccagactctctccgggaagcatctctgtgcagcgccttgcctcccgagagagc 680
 QY 601 CTCAGCCGAGACTCTCTCCGGGAGCATCTCTGTGACCCCGCTCCCTCGGAGAGAGC 660
 Db 681 caggaacactctccgagacatcatctcttcttgcgctgcagctgcagctgtgtctct 740
 QY 661 CAGGACACTCTCCGAGATCATCTCTTCTTCTGCGCTGACGCTGAGCTGCTGTCTCT 720
 Db 741 tccgtgttcttctccagcgtccgttctctgtgttgaagggcgagaaagaacatcc 800
 QY 721 TCCGTGTCTTCTCTCCAGCTCCCTTCTCTGTTGTTAAAGGGGCGAGAAAGAACTCC 780
 Db 801 tctatatactcaaaacacatttgcagacagcagtaacacacacacacacaggaagct 860
 QY 781 TGTATATTTCAACACACATTATGAGACCGATCAAACTCTCAAGAGAGAGATGGCT 840
 Db 861 gtagctgcagatttccagaaagaagaagaaggaatgtgaactgtgaatgtgaatgtca 920
 QY 841 GTAGCTGCCGATTTCCAGAGAAAGAAAGAGATGTGAATGTGAATGTGAATGTCAAT 900
 Db 921 agggctgttgcagcttcttgcagaaagaagaagaatgtgaactgtgaatgtgaatgtca 980
 QY 901 AGGGCTGTGGGACTTTCTGAAAAGAAAGCAAGAAATATGATCATCCGCTATACAGC 960
 Db 981 ttcaaaagcaagaacacacatctctacataacacagagattccccaacacacagctctt 1040
 QY 961 TTTCAAAAGCAAGAACACATCTCTCATATATCCAGAGATTCGCCAACACAGCTTCTT 1020
 Db 1041 tctaaatgcagagatgtgcctttaaataatgcagacatttcttcttgcagcag 1100
 QY 1021 TCTAAATGCCAATGATGTTGGCTTTAAATAATGACACACTTTTCTTTTTCAGACAG 1080
 Db 1101 gttcactcctgtcaacccagagctgtgagctgtgcagacacacagctctgcagccttg 1160
 QY 1081 GTCTCACCTCTACCCAGGCTGAGTGTGAGGACACCATGCTCTCTGACACCTTGG 1140
 Db 1161 acccttgcagacacagagatcctctcctcagctcctcgtgtagctgcaggaactcaagg 1220
 QY 1141 ACCTCTGGGAGCTCAAGATGATCTCTCTGCTCAGTCTCTT -AGTAGCTGGAATCTCAAGG 1199
 Db 1221 aaaggccac 1279
 QY 1200 AAGGCCACACACACTGATCACTTTTGTGTTTGTGTTGTAAGAGATTTCCGC 1259
 Db 1280 atgtgtatcagagctgtgtctcaaacctcaggtcacttgcgctcccaagatgtcggat 1339
 QY 1260 ATGTGTACAGGCTGTCTCAAACTCTAGGTTCACTTTGGCTCCCAAAAGTGTGGAT 1319

Db 1340 tacagacatgaacttgcagcggccggcccaaaataatgcacacacacacacacacag 1399
 QY 1320 TACAGACATGAACGTCCAGGCCGCCCAAAATATGACACACTTTTAAAGAGACAGACAG 1379
 Db 1400 atgagagacagagctgtgtatataaaaaaaaaaaaaa 1435
 QY 1380 ATGAGGACAGAGCTGTGTATATAAAAAAAAAAAAAA 1415
 RESULT 3
 ID Q86126 standard; cDNA: 838 BP.
 AC Q86126;
 DT 16-OCT-1995 (first entry)
 DE H4-1BB receptor protein cDNA.
 KW H4-1BB; receptor protein; immunosuppressive; autoimmune disease;
 OS organ transplantation; cell membrane ligand; ss.
 FH Homo sapiens.
 FT Key Location/Qualifiers
 FT CDS 41..808
 FT /tag- a
 FT WO9507984-A.
 PD 23-MAR-1995.
 PR 15-SEP-1994; U10457.
 PR 16-SEP-1993; US-122796.
 PA (INDV) UNIV INDIANA FOUND.
 PI KWON BS;
 DR WPI: 95-131352/17.
 DR P-PSDB: R70977.
 PT Novel cDNA encoding human receptor protein H4-1BB - useful to
 PT produce the protein which is used to treat auto-immune disease
 PS Claim 2; Fig.2; 36pp; English.
 CC Human peripheral blood lymphocyte-derived cDNA was amplified by PCR
 CC using probes based on the mouse receptor protein 4-1BB gene. The PCR
 CC product was used to screen a cDNA library of activated human T-cells.
 CC The isolated cDNA (Q86126), deposited as NRRL B21131, encoded the
 CC human homology, H4-1BB (R70977), of 4-1BB.
 SQ Sequence 838 BP; 218 A; 191 C; 215 G; 214 T;
 Query Match 59.1%; Score 836; DB 14; Length 838;
 Best Local Similarity 99.9%; Pred. No. 0.00e+00;
 Matches 837; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 1 aataagcttctgtatcatatcactgtgcagagatttcatatgggaacacagctgttaa 60
 QY 80 AATCAGCTTCTGTATATCATACCTGTGTCAGATTTCTATCATGGAACACCTGTTACAA 139
 Db 61 catagtagcacactgttgcctgtcccaacttgcagagacacacacacacacacacacacac 120
 QY 140 CATAGTAGCCACTGCTGTCGTGCTGCTCACTTTGAGAGGACAGATATATGACAGATTC 199
 Db 121 ttgtagtaactgcagcagctgtgtacattctgtgataataacaggaatcagattgcagtc 180
 QY 200 TTGTAGTAACTGCGCCAGCTGTGATCTGTGATATATACAGGAATCAGATTTGCAGTCC 259
 Db 181 ctgtccctccaatagtttctccagcagcaggtgcagaaagacacagctgtgcacatgcagga 240
 QY 260 CTGTCTCTCAAAATAGTTTCTCCAGGCGAGGTGACAAAGGACCTGTGACATATCCAGCA 319
 Db 241 gtgtaaagtttctcaagacaggaagaggttcccccacacacacacacacacacacacacac 300
 QY 380 CTGCACTCCAGGTTTCACTGCTGCGGGGCGAGATGACGATGTGTAAACGAGATTGTA 439
 Db 361 acaaggttaaaactgtgaaaaaaaggtgtgaagacagctgtgtgtgtgtgtgtgtgtgtgt 420
 QY 440 ACAAGGTCAAAAGCTGACAAAAAAGGTTGTAAAGACGTGTGTGTGTGTGTGTGTGTGTGT 499
 Db 421 tcagaacagtgcatctgtcagacctgtgacaacacgttcttgcagtggaagagctgtgtct 480

QY 500 TCAGAAAGTGGACATCTTCGACCTGACAAACGTCTTGGATGGAAAGTGTGTGCT 559
DB 481 tctgaatggagcaagagagagagagctgtctgtggaacatctccagctgacctctcc 540
QY 560 TGTGAATGGAGCAAGAGAGAGAGAGTGTGTGTGACCATCTCCAGCGACTCTCTCC 619
DB 541 gggagacatccctgtgagcccgccctgcccctgagagagagagagagagagagagag 600
QY 620 GGGAGACATCTCTGTGACCCCGCTGCCCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 679
DB 601 catctcctctctctctgtgagctgagctgagctgagctgagctgagctgagctgag 660
QY 680 CATCTCCTCTCTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 739
DB 661 gctccgcttctctgtgtttaaaggggagagagagagagagagagagagagagagag 720
QY 740 GCTCCGTTCTCTGTGTAAACGGGGCAGAAAGAACTCCGTATATATCAACCAAC 799
DB 721 attatgagacagatgagacagagagagagagagagagagagagagagagagagag 780
QY 800 ATTATGAGACAGATGACAACTACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 859
DB 781 agaagaagaagaagatgtgaactgtgaatgtgaatgtgaatgtgaatgtgaatgtga 838
QY 860 AGAAGAAGAAGAGAGATGTGAACGTGAATGAATGAATGAATGAATGAATGAATGA 917

RESULT 4
ID T39546 standard; cDNA to mRNA; 838 BP.
AC T39546;
DT 12-DEC-1996 (first entry)
DE Human receptor H4-1BB cDNA.
KW Receptor 4-1BB; H4-1BB; monoclonal antibody; T-lymphocyte; T-cell;
KW Immunostimulant; cancer; autoimmune disease; graft rejection;
KW therapy; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 41..808
FT /tag- a
FT mat_peptide 41..805
FT /tag- b
PN MO9629348-A1.
PD 26-SEP-1996.
PF 22-MAR-1996; U03965.
PR 23-MAR-1995; US-409851.
PI (INDV) UNIV INDIANA FOUND.
PA Kang C, Kwon BS;
PI MPI: 96-443138/44.
P-SDS: W04174.
F1 Monoclonal antibody specific for human receptor protein 4-1BB - used
PT to enhance proliferation and activation of T-cells for treatment of
PT cancer and to inhibit specific ligand binding for treating
PT auto-immune diseases
PS Disclosure: Page 36-37: 48pp; English.
CC A cDNA clone (T39546) codes for novel human receptor protein
CC H4-1BB (W04174), a protein that has the potential to function as
CC an accessory signaling molecule during T-cell activation and
CC proliferation. The cDNA clone was isolated from a lambda gtl1
CC cDNA library of activated human T lymphocytes by screening with a
CC PCR product obtd. by amplification of lymphocyte cDNA using
CC primers (see also T39542-45) based on the murine 4-1BB cDNA
CC (T39541). It can be used to produce recombinant H4-1BB useful
CC for isolating H4-1BB ligands, for stimulating proliferation of
CC B-cells expressing H4-1BB ligands, for blocking H4-1BB ligand
CC binding and for raising anti-H4-1BB monoclonal antibody.
SQ Sequence 838 BP; 218 A; 192 C; 214 G; 214 T;

Query Match 58.9%; Score 834; DB 24; Length 838;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;
Matches 836; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 1 atcagcttctgtatcatcacccgtgtccagatctcatcatcattggaagacagctgttaca 60

QY 80 AATCAGCTTCTAGTATCATACCTGTGCCGATTTTCATCATATGGAAACACCTTTACAA 139
DB 61 catgtagcaactgttctgtgtcctcctaacttgaagagagagagagagagagagagag 120
QY 140 CATAGTACCCACTCTGTGTGCTGTCTCAACTTACATTGAGAGGACAAGATCATGTGCGATCC 199
DB 121 ttgtagtaaacgcccagctgtgtatcatctgtgtatataaagagagagagagagag 180
QY 200 TTGTATGTAAGTCTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 259
DB 181 ctgtcctccaaatagtttctcctcagcgaggtgagcaaaagagagagagagagagagag 240
QY 260 CTGTCTCTCAATATGTTTCTCTCAGCGAGGTGAGCAAAAGACCTGTGACATATTCAGAGCA 319
DB 241 gttgaagagtttctcagagacagagagagagagagagagagagagagagagagag 300
QY 320 GTGTAAAGTGTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 379
DB 301 ctgagacccagaggttctcctcctgaggggagagagagagagagagagagagagag 360
QY 380 CTGCACTCCAGGATTTCATCTGCTGGGGCAGATGCAATGTGTGACAGATTTGTAA 439
DB 361 acaaggtcaagaactgagcaaaaaaaggttgaagagagagagagagagagagagagag 420
QY 440 ACAAGGTCAAGACTGACCAAAAAAGGTTTAAAGACTGTTGCTTTGGACATTTAAGCA 499
DB 421 tcagaagacgtgcatctgtgcagccctgcacaaactgttcttgaatggaagtgtgtct 480
QY 500 TCAGAAAGTGGACATCTGTGACCCCTGGACAACTGTTCTTGGATGGAAGTGTGTGCT 559
DB 481 tgtgaatggagcaag 540
QY 560 TGTGAATGGAGCAAG 619
DB 541 gggagacatccctgtgagcccgccctgcccctgagagagagagagagagagagagag 600
QY 620 GGGAGACATCTCTGTGACCCCGCTGCCCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 679
DB 601 catctcctctctctgtgagctgagctgagctgagctgagctgagctgagctgagct 660
QY 680 CATCTCCTCTCTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 739
DB 661 gctccgcttctctgtgtttaaaggggagagagagagagagagagagagagagagag 720
QY 740 GCTCCGTTCTCTGTGTAAACGGGGCAGAAAGAACTCCTGTATATATCAACCAAC 799
DB 721 attatgagacagatgagacagagagagagagagagagagagagagagagagagag 780
QY 800 ATTATGAGACAGATGACAACTACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 859
DB 781 agaagaagaagaagatgtgaactgtgaatgtgaatgtgaatgtgaatgtgaatgtga 838
QY 860 AGAAGAAGAAGAGAGATGTGAACGTGAATGAATGAATGAATGAATGAATGAATGA 917

RESULT 5
ID T39541 standard; cDNA to mRNA; 2347 BP.
AC T39541;
DT 12-DEC-1996 (first entry)
DE Mouse receptor 4-1BB cDNA.
KW Receptor 4-1BB; H4-1BB; monoclonal antibody; T-lymphocyte; T-cell;
KW Immunostimulant; cancer; autoimmune disease; graft rejection;
KW therapy; ss.
OS Mus sp.
FH Key Location/Qualifiers
FT CDS 146..916
FT /tag- a
FT mat_peptide 146..913
FT /tag- b
PN MO9629348-A1.
PD 26-SEP-1996.
PF 22-MAR-1996; U03965.

FT /*tag- k
FT /note- "exon 4"
FT misc_feature 2725
FT /*tag- l
FT /note- "known polymorphic site"
FT intron 2789..3328
FT /*tag- m
FT /note- "intron 4"
FT misc_feature 3063..3075
FT /*tag- n
FT /note- "indefinite interval within intron 4"
FT exon 3329..3406
FT /*tag- o
FT /note- "exon 5"
FT intron 3407..3813
FT /*tag- p
FT /note- "intron 5"
FT misc_feature 3598..3610
FT /*tag- q
FT /note- "indefinite interval within intron 5"
FT misc_feature 3653
FT /*tag- r
FT /note- "known polymorphic site"
FT exon 3814..3902
FT /*tag- s
FT /note- "exon 6"
FT intron 3903..4224
FT /*tag- t
FT /note- "intron 6"
FT misc_feature 4076..4088
FT /*tag- u
FT /note- "indefinite interval within intron 6"
FT exon 4225..4364
FT /*tag- v
FT /note- "exon 7"
FT intron 4365..6571
FT /*tag- w
FT /note- "intron 7"
FT misc_feature 4391..4392
FT /*tag- x
FT /note- "known polymorphic site"
FT misc_feature 4602..4614
FT /*tag- y
FT /note- "indefinite interval within intron 7"
FT misc_feature 6538
FT /*tag- z
FT /note- "known polymorphic site"
FT exon 6572..6677
FT /*tag- aa
FT /note- "exon 8"
FT intron 6678..9163
FT /*tag- ab
FT /note- "intron 8"
FT misc_feature 6823
FT /*tag- ac
FT /note- "known polymorphic site"
FT mutation 9106
FT /*tag- ad
FT /note- "site of 1 nucleotide deletion at known polymorphic site"
FT exon 9163..9208
FT /*tag- ae
FT /note- "exon 9"
FT misc_feature 9206
FT /*tag- af
FT /note- "known polymorphic site"
FT intron 9209..10529
FT /*tag- ag
FT /note- "intron 9"
FT misc_feature 9375
FT /*tag- ah
FT /note- "known polymorphic site"
FT exon 10530..10606

FT /*tag- ai
FT /note- "exon 10"
FT intron 10607..11596
FT /*tag- aj
FT /note- "intron 10"
FT misc_feature 11383..11395
FT /*tag- ak
FT /note- "indefinite interval within intron 10"
FT exon 11597..15022
FT /*tag- al
FT /note- "exon 11"
FT misc_feature 11907
FT /*tag- am
FT /note- "known polymorphic site"
FT misc_feature 11993
FT /*tag- an
FT /note- "known polymorphic site"
FT misc_feature 12951
FT /*tag- ao
FT /note- "known polymorphic site"
FT misc_feature 13003
FT /*tag- ap
FT /note- "known polymorphic site"
FT misc_feature 13008
FT /*tag- aq
FT /note- "known polymorphic site"
FT misc_feature 13047
FT /*tag- ar
FT /note- "known polymorphic site"
FT misc_feature 13237
FT /*tag- as
FT /note- "known polymorphic site"
FT misc_feature 13447
FT /*tag- at
FT /note- "known polymorphic site"
FT misc_feature 13538
FT /*tag- au
FT /note- "known polymorphic site"
FT misc_feature 13950
FT /*tag- av
FT /note- "known polymorphic site"
FT misc_feature 14040
FT /*tag- aw
FT /note- "known polymorphic site"
FT misc_feature 14045
FT /*tag- ax
FT /note- "known polymorphic site"
FT misc_feature 14474
FT /*tag- ay
FT /note- "known polymorphic site"
FT misc_feature 14873
FT /*tag- az
FT /note- "known polymorphic site"
FT misc_feature 14890
FT /*tag- ba
FT /note- "known polymorphic site"
FT misc_feature 14965
FT /*tag- bb
FT /note- "known polymorphic site"
FT intron 15023..15423
FT /*tag- bc
FT /note- "intron 11"
FT misc_feature 15283
FT /*tag- bd
FT /note- "known polymorphic site"
FT exon 15424..15510
FT /*tag- be
FT /note- "exon 12"
FT intron 15511..15951
FT /*tag- bf
FT /note- "intron 12"
FT misc_feature 15646..15658
FT /*tag- bg

FT /note= "indefinite interval within intron 12"
FT exon 15952..16125
FT /*tag= bh
FT /note= "exon 13"
FT misc_feature 16076
FT /*tag= bi
FT /note= "known polymorphic site"
FT intron 16126..16564
FT /*tag= bj
FT /note= "intron 13"
FT misc_feature 16242
FT /*tag= bk
FT /note= "known polymorphic site"
FT misc_feature 16369..16381

...
Note: remainder of annotations omitted.

Query Match 7.4%; Score 105; DB 21; Length 24025;
Best Local Similarity 76.1%; Pred. No. 1,93e-42;
Matches 175; Conservative 0; Mismatches 52; Indels 3; Gaps 3;

7404 ggaagttcaagaccagcctcaccacatgagaaacccatctactataaaaaaaaaa 7463
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
1286 GGAGTTTGAGACGACGCTGTACAACTGGCAATGCGAATGCTTTACCAACAAACAA 1227
7464 aaaaattagccggggtgtgcttaatcgttaaccacactcagagagctgagc 7523
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1226 AAAAGTTAATCATGAGTGTGTGGCCCTCTTGTAGTTCAGACTAC-AGGAGACTAGGC 1168
Db 7524 aggaagaatcgttgaaac-ccaggaagcagaaggttcagtgagcacaagatcgaccatlgc 7582
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
1167 AGGAGATCATCTTGAGCTCCAGAGGTCAAGGCTGCAGAGAGCCAGTGTGTCCTGC 1108
CP 7583 actccagcctagcacaagaagtgaatccatctcaaaaaaaaaa 7632
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
CP 1107 ACTCCAGCCTGGGTGACA-GAGTGAGACCTGTCAAAAAAAAAAAAAA 1059

RESULT 13

ID T17455 standard; cDNA; 24025 BP.

AC T17455;

DT 07-OCT-1996 (first entry)

DE Mutated BRCA1 genomic sequence from sample set MSKCC family 19921.

KW Cancer therapy; breast and ovarian cancer predisposing gene; immunogen;

KW antibody production; germ-line alteration; probe; lesion neoplasia; human;

KW gene therapy; protein replacement therapy; protein mimetic; BRCA1; ds.

OS Homo sapiens.

FH Key Location/Qualifiers

FT exon 256..355

FT /*tag= a

FT /note= "exon 1"

FT intron 356..1512

FT /*tag= b

FT /note= "intron 1"

FT misc_feature 1295

FT /*tag= c

FT /note= "known polymorphic site"

FT exon 1513..1611

FT /*tag= d

FT /note= "exon 2"

FT intron 1612..2206

FT /*tag= e

FT /note= "intron 2"

FT misc_feature 1925..1937

FT /*tag= f

FT /note= "indefinite interval within intron 2"

FT misc_feature 2141

FT /*tag= g

FT /note= "known polymorphic site"

FT exon 2207..2260

FT /*tag= h

FT /note= "exon 3"

FT intron 2261..2677

FT /*tag= i
FT /note= "intron 3"
FT misc_feature 2569..2581
FT /*tag= j
FT /note= "indefinite interval within intron 3"
FT exon 2678..2788
FT /*tag= k
FT /note= "exon 4"
FT misc_feature 2725
FT /*tag= l
FT /note= "known polymorphic site"
FT intron 2789..3328
FT /*tag= m
FT /note= "intron 4"
FT misc_feature 3063..3075
FT /*tag= n
FT /note= "indefinite interval within intron 4"
FT exon 3329..3406
FT /*tag= o
FT /note= "exon 5"
FT intron 3407..3813
FT /*tag= p
FT /note= "intron 5"
FT misc_feature 3598..3610
FT /*tag= q
FT /note= "indefinite interval within intron 5"
FT misc_feature 3653
FT /*tag= r
FT /note= "known polymorphic site"
FT exon 3814..3902
FT /*tag= s
FT /note= "exon 6"
FT intron 3903..4224
FT /*tag= t
FT /note= "intron 6"
FT mutation 4223
FT /*tag= u
FT /note= "site of 1 nucleotide deletion"
FT misc_feature 4076..4088
FT /*tag= v
FT /note= "indefinite interval within intron 6"
FT exon 4225..4364
FT /*tag= w
FT /note= "exon 7"
FT intron 4365..6571
FT /*tag= x
FT /note= "intron 7"
FT misc_feature 4391..4392
FT /*tag= y
FT /note= "known polymorphic site"
FT misc_feature 4602..4614
FT /*tag= z
FT /note= "indefinite interval within intron 7"
FT misc_feature 6538
FT /*tag= aa
FT /note= "known polymorphic site"
FT exon 6572..6677
FT /*tag= ab
FT /note= "exon 8"
FT intron 6678..9163
FT /*tag= ac
FT /note= "intron 8"
FT misc_feature 6823
FT /*tag= ad
FT /note= "known polymorphic site"
FT misc_feature 9106
FT /*tag= ae
FT /note= "known polymorphic site"
FT exon 9164..9209
FT /*tag= af
FT /note= "exon 9"
FT misc_feature 9207
FT /*tag= ag


```

FT /note- "known polymorphic site"
FT intron 9210..10530
FT /*tag- ah
FT /note- "intron 9"
FT misc_feature 9376
FT /*tag- ai
FT /note- "known polymorphic site"
FT exon 10531..10607
FT /*tag- aj
FT /note- "exon 10"
FT intron 10608..11597
FT /*tag- ak
FT /note- "intron 10"
FT misc_feature 11384..11396
FT /*tag- al
FT /note- "indefinite interval within intron 10"
FT exon 11598..15023
FT /*tag- am
FT /note- "exon 11"
FT misc_feature 11908
FT /*tag- an
FT /note- "known polymorphic site"
FT misc_feature 11994
FT /*tag- ao
FT /note- "known polymorphic site"
FT misc_feature 12952
FT /*tag- ap
FT /note- "known polymorphic site"
FT misc_feature 13004
FT /*tag- aq
FT /note- "known polymorphic site"
FT misc_feature 13009
FT /*tag- ar
FT /note- "known polymorphic site"
FT misc_feature 13048
FT /*tag- as
FT /note- "known polymorphic site"
FT misc_feature 13238
FT /*tag- at
FT /note- "known polymorphic site"
FT misc_feature 13448
FT /*tag- au
FT /note- "known polymorphic site"
FT misc_feature 13539
FT /*tag- av
FT /note- "known polymorphic site"
FT misc_feature 13951
FT /*tag- aw
FT /note- "known polymorphic site"
FT misc_feature 14041
FT /*tag- ax
FT /note- "known polymorphic site"
FT misc_feature 14046
FT /*tag- ay
FT /note- "known polymorphic site"
FT misc_feature 14475
FT /*tag- az
FT /note- "known polymorphic site"
FT misc_feature 14874
FT /*tag- ba
FT /note- "known polymorphic site"
FT misc_feature 14891
FT /*tag- bb
FT /note- "known polymorphic site"
FT misc_feature 14966
FT /*tag- bc
FT /note- "known polymorphic site"
FT intron 15024..15424
FT /*tag- bd
FT /note- "intron 11"
FT misc_feature 15284
FT /*tag- be
FT /note- "known polymorphic site"

```

```

FT exon 15425..15511
FT /*tag- bf
FT /note- "exon 12"
FT intron 15512..15952
FT /*tag- bg
FT /note- "intron 12"
FT misc_feature 15647..15659
FT /*tag- bh
FT /note- "indefinite interval within intron 12"
FT exon 15953..16126
FT /*tag- bi
FT /note- "exon 13"
FT misc_feature 16077
FT /*tag- bj
FT /note- "known polymorphic site"
FT intron 16127..16565
FT /*tag- bk
FT /note- "intron 13"
FT misc_feature 16243
FT /*tag- bl

```

...
Note: remainder of annotations omitted.

Query Match 7.4%; Score 105; DB 21; Length 24025;
Best Local Similarity 76.1%; Pred. No. 1,93e+42;
Matches 175; Conservative 0; Mismatches 52; Indels 3; Gaps 3;

```

Db 7403 ggaattcaagaccagcctgcaacatgagagaaccccatctactaataaaaaa 7462
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Cp 1286 GGAGTTGAGACCGACTGACACATGCGGAATGCACTTTACCAACAAAAACAA 1227
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7463 aaaaattagccgggtggtgttatacctgtatccactactcaaggagctgagc 7522
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 1226 AAGAATTAGTACAGTGTGCGGCCCTCTGTGTAGTCCAGCTACT AGAGAAGTGAAGC 1168
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7523 agagaatgccttgaaac-ccaaggaagcagaaggttgcaagtaagccaagatgcacattgc 7581
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 1167 AGGAGATCACTTGAGCTCCAGCAAGCTGCAAGAGCGCATGCTGTGCTGCACTGC 1108
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7582 actccagccttagccaacaagaagtgaactcatctcacaataaaaaa 7631
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 1107 ACTCCAGCCTGGTGACCA-GAGTGAGACCTGTCAAAAAAAAAAAAAA 1059
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 14
ID T17516 standard: cDNA: 24026 BP.
AC T17516;
DT 04-Oct-1996 (first entry)
DE Mutated BRCA1 genomic sequence from PM16.
KW Cancer therapy; breast and ovarian cancer predisposing gene; immunogen;
KW antibody production; germline alteration; probe; lesion neoplasia; human;
KW gene therapy; protein replacement therapy; protein mimetic; BRCA1; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT exon 256..355
FT /*tag- a
FT /note- "exon 1"
FT intron 356..1512
FT /*tag- b
FT /note- "intron 1"
FT misc_feature 1295
FT /*tag- c
FT /note- "known polymorphic site"
FT exon 1513..1611
FT /*tag- d
FT /note- "exon 2"
FT intron 1612..2206
FT /*tag- e
FT /note- "intron 2"
FT misc_feature 1925..1937
FT /*tag- f
FT /note- "indefinite interval within intron 2"
FT misc_feature 2141

```

```
FT /tag- g
FT /note- "known polymorphic site"
FT exon 2207..2260
FT /tag- h
FT /note- "exon 3"
FT intron 2261..2677
FT /tag- i
FT /note- "intron 3"
FT misc.feature 2569..2584
FT /tag- j
FT /note- "indefinite interval within intron 3"
FT exon 2678..2788
FT /tag- k
FT /note- "exon 4"
FT misc.feature 2725
FT /tag- l
FT /note- "known polymorphic site"
FT intron 2789..3328
FT /tag- m
FT /note- "intron 4"
FT misc.feature 3063..3075
FT /tag- n
FT /note- "indefinite interval within intron 4"
FT exon 3329..3406
FT /tag- o
FT /note- "exon 5"
FT intron 3407..3813
FT /tag- p
FT /note- "intron 5"
FT misc.feature 3598..3610
FT /tag- q
FT /note- "indefinite interval within intron 5"
FT misc.feature 3653
FT /tag- r
FT /note- "known polymorphic site"
FT exon 3814..3902
FT /tag- s
FT /note- "exon 6"
FT intron 3903..4224
FT /tag- t
FT /note- "intron 6"
FT misc.feature 4076..4088
FT /tag- u
FT /note- "indefinite interval within intron 6"
FT exon 4225..4364
FT /tag- v
FT /note- "exon 7"
FT intron 4365..6571
FT /tag- w
FT /note- "intron 7"
FT misc.feature 4391..4392
FT /tag- x
FT /note- "known polymorphic site"
FT misc.feature 4602..4614
FT /tag- y
FT /note- "indefinite interval within intron 7"
FT misc.feature 6538
FT /tag- z
FT /note- "known polymorphic site"
FT exon 6572..6677
FT /tag- aa
FT /note- "exon 8"
FT intron 6678..9163
FT /tag- ab
FT /note- "intron 8"
FT misc.feature 6823
FT /tag- ac
FT /note- "known polymorphic site"
FT misc.feature 9106
FT /tag- ad
FT /note- "known polymorphic site"
FT exon 9164..9209
FT /tag- ae

FT /note- "exon 9"
FT misc.feature 9207
FT /tag- af
FT /note- "known polymorphic site"
FT intron 9210..10530
FT /tag- ag
FT /note- "intron 9"
FT misc.feature 9376
FT /tag- ah
FT /note- "known polymorphic site"
FT exon 10531..10607
FT /tag- ai
FT /note- "exon 10"
FT intron 10608..11597
FT /tag- aj
FT /note- "intron 10"
FT misc.feature 11384..11396
FT /tag- ak
FT /note- "indefinite interval within intron 10"
FT exon 11598..15023
FT /tag- al
FT /note- "exon 11"
FT misc.feature 11908
FT /tag- am
FT /note- "known polymorphic site"
FT misc.feature 11994
FT /tag- an
FT /note- "known polymorphic site"
FT misc.feature 12952
FT /tag- ao
FT /note- "known polymorphic site"
FT misc.feature 13004
FT /tag- ap
FT /note- "known polymorphic site"
FT misc.feature 13009
FT /tag- aq
FT /note- "known polymorphic site"
FT misc.feature 13048
FT /tag- ar
FT /note- "known polymorphic site"
FT misc.feature 13238
FT /tag- as
FT /note- "known polymorphic site"
FT misc.feature 13448
FT /tag- at
FT /note- "known polymorphic site"
FT misc.feature 13539
FT /tag- au
FT /note- "known polymorphic site"
FT misc.feature 13951
FT /tag- av
FT /note- "known polymorphic site"
FT misc.feature 14041
FT /tag- aw
FT /note- "known polymorphic site"
FT misc.feature 14046
FT /tag- ax
FT /note- "known polymorphic site"
FT misc.feature 14475
FT /tag- ay
FT /note- "known polymorphic site"
FT misc.feature 14874
FT /tag- az
FT /note- "known polymorphic site"
FT misc.feature 14891
FT /tag- ba
FT /note- "known polymorphic site"
FT misc.feature 14966
FT /tag- bb
FT /note- "known polymorphic site"
FT intron 15024..15424
FT /tag- bc
FT /note- "intron 11"
```



```

FT /note- "known polymorphic site"
FT misc_feature 9106
FT /tag- ad
FT /note- "known polymorphic site"
FT exon 9164..9209
FT /tag- ae
FT /note- "exon 9"
FT misc_feature 9207
FT /tag- af
FT /note- "known polymorphic site"
FT intron 9210..10530
FT /tag- ag
FT /note- "intron 9"
FT misc_feature 9376
FT /tag- ah
FT /note- "known polymorphic site"
FT exon 10531..10607
FT /tag- ai
FT /note- "exon 10"
FT intron 10608..11597
FT /tag- aj
FT /note- "intron 10"
FT misc_feature 11384..11396
FT /tag- ak
FT /note- "indefinite interval within intron 10"
FT exon 11598..115023
FT /tag- al
FT /note- "exon 11"
FT misc_feature 11908
FT /tag- am
FT /note- "known polymorphic site"
FT misc_feature 11994
FT /tag- an
FT /note- "known polymorphic site"
FT misc_feature 12952
FT /tag- ao
FT /note- "known polymorphic site"
FT misc_feature 13004
FT /tag- ap
FT /note- "known polymorphic site"
FT /tag- aq
FT /note- "known polymorphic site"
FT misc_feature 13048
FT /tag- ar
FT /note- "known polymorphic site"
FT misc_feature 13238
FT /tag- as
FT /note- "known polymorphic site"
FT misc_feature 13448
FT /tag- at
FT /note- "known polymorphic site"
FT misc_feature 13539
FT /tag- au
FT /note- "known polymorphic site"
FT misc_feature 13951
FT /tag- av
FT /note- "known polymorphic site"
FT misc_feature 14041
FT /tag- aw
FT /note- "known polymorphic site"
FT misc_feature 14046
FT /tag- ax
FT /note- "known polymorphic site"
FT misc_feature 14475
FT /tag- ay
FT /note- "known polymorphic site"
FT misc_feature 14874
FT /tag- az
FT /note- "known polymorphic site"
FT misc_feature 14891
FT /tag- ba
FT /note- "known polymorphic site"

```

```

FT misc_feature 14966
FT /tag- bb
FT /note- "known polymorphic site"
FT intron 15024..15424
FT /tag- bc
FT /note- "intron 11"
FT misc_feature 15284
FT /tag- bd
FT /note- "known polymorphic site"
FT exon 15425..15511
FT /tag- be
FT /note- "exon 12"
FT intron 15512..15952
FT /tag- bf
FT /note- "intron 12"
FT misc_feature 15647..15659
FT /tag- bg
FT /note- "indefinite interval within intron 12"
FT exon 15953..16126
FT /tag- bh
FT /note- "exon 13"
FT misc_feature 16077
FT /tag- bi
FT /note- "known polymorphic site"
FT intron 16127..16565
FT /tag- bj
FT /note- "intron 13"
FT misc_feature 16243
FT /tag- bk
FT /note- "known polymorphic site"
FT misc_feature 16370..16382
FT /tag- bl

```

Note: remainder of annotations omitted.

Query Match 7.4%; Score 105; DB 21; Length 24026;
 Best Local Similarity 76.1%; Pred. No. 1,93e-42;
 Matches 175; Conservative 0; Mismatches 52; Indels 3; Gaps 3;

```

Db 7404 ggaagtcagaagcagcctgacccaacatgagaaaccccatctactataaaaaa 7463
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 1286 GGAATTGAGACGACGCTGTACAAACATGCGCAAAATGCCATCTTACCAACAAAAACAA 1227
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7464 aaaaatagccggggtgtgtgtatgcctgtatcccgctactcagaggtgagc 7523
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 1226 AAAGTTAGTCAGGTGTGTGCGCCCTCTTGTAGTTCAGTACT-AGGAGACTGAGGC 1168
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7524 agagaatcgctgaac-ccaggaagcagaggttgcagtgagccaagatgcaccattgc 7582
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 1167 AGGAGATCACTGAGCTCCAGAGGTCAAGGCTGCAGAGAGCCATGCTGTGCTCACTGC 1108
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7583 actcagcctaggcaacaagatgaaactccatctcaaaaaa 7632
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 1107 ACTCCAGCCTGGTGNACA-GAGTGAGACCTGTCAAAAAA 1059
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: Wed Mar 25 13:32:37 1998
 Job time : 213 secs.

Oy	61	TCDICRCKGKGFRTKRKESSTSNMBCDCTPGFHGIGAGCSMCEDDKQGGQLTRKGGKDC	120
Dd	121	cfcgtndckrgtcrpwctcsldgksvlyngtkerdvcvgspadlspgassvtppapare	180
Oy	121	CFGTNDCKRGICRPWTCSLDGKSVLVNGTKERDVYCGSPADLSPASSVTTPAPARE	180
Dd	181	pghspqisffaltsstallflfllrtlrsvkrgrkkllylfkqpmrpyvtctgeedg	240
Oy	181	PGHSPQISFFALTSALLFLFLLRLRSVKRKKLILYIFKPMPRVOTTEBDG	240
Dd	241	cscrfeeeeggccl 255 ~	
Oy	241	CSCRFEDEEGGCCL 255	
RESULT	2		
ENTRY	B32393	#type complete	
TITLE	T-cell antigen 4-1BB precursor - mouse		
ORGANISM	#formal_name Mus musculus #common_name house mouse		
DATE	17-Jul-1992	#sequence_revision 17-Jul-1992	#text_change
SESSIONS	02-Aug-1996		
REFERENCE	B32393; I48879		
#authors	A32393		
#journal	Kwon, B.S.; Weissman, S.M.		
#title	Proc. Natl. Acad. Sci. U.S.A. (1989) 86:1963-1967		
#cross-references MUID	CDNA sequence of two inducible T-cell genes.		
#accession	B32393		
#molecule_type	mRNA		
#residues	1-256	#label KMO	
#cross-references GB:	J04492		
REFERENCE	I48879		
#authors	Kwon, B.S.; Kozak, C.A.; Kim, K.K.; Pickard, R.T.		
#journal	J. Immunol. (1994) 152:2256-2262		
#title	Genomic Organization and Chromosomal Localization of the T-cell Antigen 4-1BB.		
#cross-references MUID	94179805		
#accession	I48879		
#status	preliminary; translated from GB/EMBL/DDBJ		
#molecule_type	DNA		
#residues	1-256	#label RES	
#cross-references EMBL:	U02567; NID:g409177; CDS_PID:g409178		
GENETICS	35/1; 69/1; 115/1; 138/2; 156/3; 181/1; 225/1		
#introns			
CLASSIFICATION	#superfamily NGF receptor repeat homology		
KEYWORDS	transmembrane protein		
FEATURE			
1-23			
94-236			
MARY			
Query Match	58.0%; Score 1116; DB 16; Length 256;		
Best Local Similarity	57.3%; Pred. No. 6-29e-154;		
Matches 145; Conservative 49; Mismatches 55; Indels 4; Gaps 4.			
Dd	1	mgnncynuvviavillvygcgvagvscnscdqpltfcrky-npyokscpsstfsfsgqp 59	
Oy	1	MGNNCYNIVATILLVLFNERTRILODPSCNPAGFCDDNNNOJCPCPPRFSSSAGQR 60	
Dd	60	nchlcrvaaygfirfkfssshnaecieiefghclgpqctcekdrcpgelrtkgqctc 119	
Oy	61	TCDICRCKGKGFRTKRKESSTSNMBCDCTPGFHGIGAGCSMCEDDKQGGQLTRKGGKDC 120	
Dd	120	slgftndnggygvcpwcnsldgrrsvlktttekdvwcvpprvsfaspstlstvtpggp 179	
Oy	121	CFGFTNDCKRGICRPWTCNSLDGKSVLVNGTKERDVYCGSPADLSGASVTPPAPAR 179	
Dd	180	g-shalsqlvltflafats-allatalfllslsvkwirkkfphikfpkkttgaageed 237	
Oy	180	EPGHPQIISFFALTSALLFLFELLRLRSVYKRGKKLILYIFKQPMRPVOTTGED 239	
Dd	238	acscrtpqeeg 250	

```

QY      240   GCSGRFPEEBEGG    252
          :|||||:-|||
RESULT      3
ENTRY       A49053 #type complete
TITLE       CD27 antigen precursor - mouse
ALTERNATE_NAMES
ORGANISM    #formal_name Mus musculus #common_name house mouse
DATE        19-Dec-1993 #sequence_revision 22-Apr-1995 #text_change
                22-Apr-1995
ACCESSIONS  A49053
REFERENCE    A49053
AUTHORS      Gravesstein, L.A.; Blom, B.; Nolten, L.A.; de Vries, E.; van
                der Horst, G.; Osendorp, F.; Borst, J.; Loenen, W.A.
#journal     Eur. J. Immunol. (1993) 23:943-950
#title       Cloning and expression of murine CD27: comparison with 4-1BBL
                another lymphocyte-specific member of the nerve growth
                factor receptor family.
#cross-references MIMD:93209296
#accession   A49053
#molecule_type mRNA
#residues    1-250 #label GRA
#cross-references NCBIN:128168; NCBIPI:128169
CLASSIFICATION
#note        sequence extracted from NCBI backbone
KEYWORDS      #superfamily CD27 antigen; NGF receptor repeat homology
                duplication; glycoprotein; homodimer; receptor; surface
                antigen; T-cell; transmembrane protein
FEATURE
1-20          #domain signal sequence #status predicted #label SIG\
21-182        #domain extracellular #status predicted #label EXT\
22-63         #domain NGF receptor repeat homology #label NG1\
65-105        #domain NGF receptor repeat homology #label NG2\
121-179       #region proline/serine/threonine-rich\
183-202       #domain transmembrane #status predicted #label TMN\
203-250       #domain intracellular #status predicted #label INT\
95,162        #binding_site carbohydrate (asn) (covalent) #status
                predicted
SUMMARY       #length 250 #molecular_weight 28164 #checksum 6216
Query Match   10.7% Score 206; DB 2; Length 250;
Best Local Similarity 37.1%; Pred. NO.3-29e-14;
Matches 26; Conservative 14; Mismatches 26; Indels 4; Gaps 4
Db            53 cegdtlaagcdpcipgtsfspdylrhphcscsrhngsgflrn-clvtanaecscsknwq 111
              |:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY            37 CDNNRNRQI-CSPC-PNPSFFSSAGGQR-CDICRCKGVFRTRKCCSSYNNAECCTCGFH 93
              ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB            112 crdgedcted 121
OY           QY 94 CLGAGCSMCE 103
RESULT      4
ENTRY       A46517 #type complete
TITLE       CD27 antigen precursor - human
ALTERNATE_NAMES
ORGANISM    #formal_name Homo sapiens #common_name man
DATE        18-Jun-1993 #sequence_revision 22-Apr-1995 #text_change
                23-Feb-1997
ACCESSIONS  A46517
REFERENCE    A46517
AUTHORS      Loenen, W.A.; Gravesstein, L.A.; Beumer, S.; Melief, C.J.;
                Hagemeijer, A.; Borst, J.
#journal     J. Immunol. (1992) 149:3937-3943
#title       Genomic organization and chromosomal localization of the
                human CD27 gene.
#cross-references MIMD:93094588
#accession   A46517
#status      not compared with conceptual translation
#molecule_type DNA
#residues    1-260 #label IOE
#cross-references NCBIPI:120386

```


##molecule_type DNA
##accession 1-349 ##label KOL
##cross-references EMBL:X67117
##experimental_source strain India-1967, isolate Ind3
CLASSIFICATION #superfamily NGF receptor repeat homology
FEATURE
68-109 #domain NGF receptor repeat homology #label NG2\
110-151 #domain NGF receptor repeat homology #label NG3
SUMMARY #length 349 #molecular-weight 38189 #checksum 2016

Query Match 9.3%; Score 179; DB 9; Length 349;
Best Local Similarity 34.1%; Pred. No. 2.04e-10;
Matches 29; Conservative 20; Mismatches 29; Indels 7; Gaps 4;

Db 44 clscppgyarldskntqctpgsqtftsmhlpclscngscnsgnscntt 103
QY 28 CSNCPAGTF---CDNNKNOICSPCPNFSFSSAGGQ-RTCDICR-CKGKVFRTREKCSST 81
Db 104 hnricecspgyyalkgssgc-kac 127
QY 82 SNABEDCTPFHCLGAGCSMCEQDC 106

RESULT 8
ENTRY B43692 #type complete
TITLE T2 protein - rabbit fibroma virus
ORGANISM #formal_name rabbit fibroma virus, Shope fibroma virus
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 23-May-1997

ACCESSIONS B43692
REFERENCE A43692
#authors Upton, C.; Delange, A.M.; McFadden, G.
#journal Virology (1987) 160:20-30
#title Tumorigenic poxviruses: genomic organization and DNA sequence of the telomeric region of the Shope fibroma virus genome.
#accession B43692
#status preliminary
#molecule_type DNA
#residues 1-325 #label UPT
#cross-references GB:M17433
CLASSIFICATION #superfamily myxoma virus T2 protein; NGF receptor repeat homology

FEATURE
64-105 #domain NGF receptor repeat homology #label NG2\
106-147 #domain NGF receptor repeat homology #label NG3\
SUMMARY #length 325 #molecular-weight 35132 #checksum 4629

Query Match 9.0%; Score 174; DB 6; Length 325;
Best Local Similarity 31.3%; Pred. No. 9.91e-10;
Matches 25; Conservative 15; Mismatches 35; Indels 5; Gaps 4;

Db 49 asrlcpgpsntvcspcdgdtftasthnapacvscrgpctghlsgsqcdtrhdyvncst 108
QY 33 AGTCDDNNRNOICSPCPNFSFSSAGGQ-RTCDICR-CKGKVFRTREKCSSTNAECDCTP 90
Db 109 gnycllkgqngcricapqtk 128
QY 91 GFHCL--G-AGCSMCEQDC 107

RESULT 9
ENTRY A46515 #type complete
TITLE B cell-associated surface molecule CD40 - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995

ACCESSIONS A46515
REFERENCE A46515
#authors Grimaldi, J.C.; Torres, R.; Kozak, C.A.; Chang, R.; Clark, E.A.; Howard, M.; Cockayne, D.A.
#journal J. Immunol. (1992) 149:3921-3926
#title Genomic structure and chromosomal mapping of the murine CD40 gene.

##cross-references MUID:93094586
#accession A46515
#status preliminary; not compared with conceptual translation
##molecule_type nucleic acid
##residues 1-289 #label GRI
##cross-references NCBI:P120357
##experimental_source BALB/c, liver
#note sequence extracted from NCBI backbone
SUMMARY #length 289 #molecular-weight 32111 #checksum 579

Query Match 8.9%; Score 171; DB 16; Length 289;
Best Local Similarity 31.2%; Pred. No. 2.54e-09;
Matches 43; Conservative 21; Mismatches 58; Indels 16; Gaps 11;

Db 51 ctalekq-chnpcdsgefsaqwreirchqhhrhcepnqglrvkkgtaesdvtctkeqg 109
QY 37 CDN-NRNOICSPCPNFSFSSA-GGQRTCDICRCK-g-VRTREKCSSTNAECDCTPFG 92
Db 110 hctskdeacaghtpcipjgfyvmematetdvtchpcpygffsngslfckypwtscd 169
QY 93 HCLGAGCSMCEQD--CKQG---QEL-T-KKG--CKDCCFGFNDQKR--GICRWTCSL 141
Db 170 knlevlqkysqtnvicg 187
QY 142 DGKSVLVNGTKERDVCVG 159

RESULT 10
ENTRY A46476 #type complete
TITLE CD40 - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 16-Feb-1997

ACCESSIONS A46476
REFERENCE A46476
#authors Torres, R.M.; Clark, E.A.
#journal J. Immunol. (1992) 148:620-626
#title Differential increase of an alternatively polyadenylated mRNA species of murine CD40 upon B lymphocyte activation.
#cross-references MUID:92105763
#accession A46476
#status preliminary
#molecule_type mRNA
##residues 1-305 #label TOR
##cross-references NCBI:N75206; NCBI:P75207
#note sequence extracted from NCBI backbone
KEYWORDS transmembrane protein
SUMMARY #length 305 #molecular-weight 33617 #checksum 5203

Query Match 8.9%; Score 171; DB 16; Length 305;
Best Local Similarity 31.2%; Pred. No. 2.54e-09;
Matches 43; Conservative 21; Mismatches 58; Indels 16; Gaps 11;

Db 51 ctalekq-chnpcdsgefsaqwreirchqhhrhcepnqglrvkkgtaesdvtctkeqg 109
QY 37 CDN-NRNOICSPCPNFSFSSA-GGQRTCDICRCK-g-VRTREKCSSTNAECDCTPFG 92
Db 110 hctskdeacaghtpcipjgfyvmematetdvtchpcpygffsngslfckypwtscd 169
QY 93 HCLGAGCSMCEQD--CKQG---QEL-T-KKG--CKDCCFGFNDQKR--GICRWTCSL 141
Db 170 knlevlqkysqtnvicg 187
QY 142 DGKSVLVNGTKERDVCVG 159

RESULT 11
ENTRY S12783 #type complete
TITLE OX40 antigen precursor - rat
ALTERNATE_NAMES nerve growth factor receptor homolog
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 30-May-1997

 WIDE (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
 Distribution rights by Intelligenetics, Inc.

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm
 Date: Tue Mar 24 17:03:43 1998; MasPar time 7.28 Seconds
 Similar output not generated. 742.674 Million cell updates/sec

Title: >US-08-236-918A-8
 Description: (1-255) From 5674704.pep
 Perfect Score: 1925
 Sequence: 1 MGNSCYNIVATLLVLNFER.....QEEGDCSCRFPEEEGCGCEL 255

Scoring table: PAM 150
 Gap 11

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot34
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11

Statistics: Mean 45.059; Variance 89.948; scale 0.501

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Rank	Score	Match	Length	DB	ID	Description	Pred. No.
1	1925	100.0	255	1	41BB_HUMAN	4-1BB LIGAND RECEPTOR	0.00e+00
2	1116	58.0	256	1	41BB_MOUSE	4-1BB LIGAND RECEPTOR	6.16e-200
3	206	10.7	250	2	CD27L_MOUSE	CD27L RECEPTOR PRECUR	1.05e-17
4	199	10.3	260	2	CD27L_HUMAN	CD27L RECEPTOR PRECUR	1.67e-16
5	183	9.5	326	10	VT2_MXVL	TUMOR NECROSIS FACTOR	8.48e-14
6	183	9.5	415	9	TNRC_MOUSE	LYMPHOTOXIN-BETA RECE	8.48e-14
7	179	9.3	349	10	VC22_VAVR	PROTEIN C22/B28 HOMOL	3.93e-13
8	174	9.0	325	10	VT2_STVKA	TUMOR NECROSIS FACTOR	2.63e-12
9	171	8.9	289	9	CD40_MOUSE	CD40L RECEPTOR PRECUR	8.18e-12
10	166	8.6	271	7	OX40_RAT	OX40L RECEPTOR PRECUR	5.33e-11
11	158	8.2	435	9	TNRC_HUMAN	LYMPHOTOXIN-BETA RECE	1.03e-09
12	140	7.3	272	7	OX40_MOUSE	OX40L RECEPTOR PRECUR	6.53e-07
13	141	7.3	277	7	OX40_HUMAN	OX40L RECEPTOR PRECUR	4.60e-07
14	138	7.2	425	6	NGFR_RAT	LOW-AFFINITY NERVE GR	1.31e-06
15	136	7.1	277	2	CD40_HUMAN	CD40L RECEPTOR PRECUR	2.63e-06
16	133	7.0	427	6	NGFR_HUMAN	LOW-AFFINITY NERVE GR	3.71e-06
17	133	6.9	461	9	TNR2_HUMAN	TUMOR NECROSIS FACTOR	7.37e-06
18	127	6.6	986	3	EPH3_HUMAN	TYROSINE-PROTEIN KINA	5.62e-05
19	127	6.6	993	3	EPH3_MOUSE	TYROSINE-PROTEIN KINA	5.62e-05
20	123	6.4	1609	6	LMG1_HUMAN	LAMININ GAMMA-1 CHAIN	2.12e-04
21	122	6.3	335	3	FASA_HUMAN	FAST RECEPTOR PRECURS	2.95e-04
22	121	6.3	474	9	TNR2_MOUSE	TUMOR NECROSIS FACTOR	4.09e-04

Rank	Score	Match	Length	DB	ID	Description	Pred. No.
23	121	6.3	1607	6	LMG1_MOUSE	LAMININ GAMMA-1 CHAIN	4.09e-04
24	119	6.2	416	6	NGFR_CHICK	LOW-AFFINITY NERVE GR	7.84e-04
25	120	6.2	577	9	TBRM_MOUSE	THROMBOMODULIN PRECUR	5.67e-04
26	120	6.2	984	3	EPH2_HUMAN	TYROSINE-PROTEIN KINA	5.67e-04
27	119	6.2	1788	6	LMB1_DROME	LAMININ BETA-1 CHAIN	7.84e-04
28	117	6.1	713	9	TS4A_GIALA	MAJOR SURFACE-LABELLED	1.50e-03
29	118	6.1	755	2	COMP_RAT	CARTRIDGE OLIGOMERIC	1.08e-03
30	118	6.1	984	3	EPH2_RAT	TYROSINE-PROTEIN KINA	1.08e-03
31	117	6.1	1680	4	FUR2_DROME	FURIN-LIKE PROTEASE 2	1.50e-03
32	117	6.1	1786	6	LMB1_HUMAN	LAMININ BETA-1 CHAIN	1.50e-03
33	115	6.0	987	5	HTK_MOUSE	TYROSINE-PROTEIN KINA	2.83e-03
34	115	6.0	995	3	EPH3_CHICK	TYROSINE-PROTEIN KINA	2.83e-03
35	113	5.9	327	3	FASA_MOUSE	FAST RECEPTOR PRECURS	5.34e-03
36	113	5.9	461	9	TNR1_RAT	TUMOR NECROSIS FACTOR	3.89e-03
37	114	5.9	987	5	HTK_HUMAN	TYROSINE-PROTEIN KINA	3.89e-03
38	114	5.9	1005	3	EHK1_RAT	TYROSINE-PROTEIN KINA	3.89e-03
39	114	5.9	1712	9	TGFB_RAT	LATENT TRANSFORMING G	3.89e-03
40	112	5.8	454	9	TNR1_MOUSE	TUMOR NECROSIS FACTOR	7.31e-03
41	108	5.6	1039	1	ATHL_HUMAN	POTASSIUM-TRANSPORTIN	2.53e-02
42	107	5.6	1786	6	LMB1_MOUSE	LAMININ BETA-1 CHAIN	3.43e-02
43	107	5.6	1801	6	LMB2_RAT	LAMININ BETA-2 CHAIN	3.43e-02
44	106	5.5	595	2	CD30L_HUMAN	CD30L RECEPTOR PRECUR	4.65e-02
45	106	5.5	915	7	PAC6_RAT	SERINE PROTEINASE PC6 P	4.65e-02

ALIGNMENTS

Result ID	1	Standard	PRT	255 AA
AC	007011			
DT	01-FEB-1995 (REL. 31, CREATED)			
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)			
DT	01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)			
DE	4-1BB LIGAND RECEPTOR PRECURSOR (T-CELL ANTIGEN 4-1BB HOMOLOG)			
DE	(T-CELL ANTIGEN IIA).			
GN	IIA.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUDARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-BLOOD:			
RX	MEDLINE; 94374434.			
RA	ALDERSON M.R., SMITH C.A., TOUGH T.W., DAVIS-SMITH T., ARMITAGE R.J.,			
RA	FAIK B., ROUX E., BAKER E., SUTHERLAND G.R., DIN W.S., GOODWIN R.G.,			
RL	EUR. J. IMMUNOL. 24:2219-2227(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-BLOOD:			
RX	MEDLINE; 94085794.			
RA	SCHWARTZ H., TUCKERELL J., LOTZ M.;			
RA	GENE 134:295-298(1993).			
CC	- FUNCTION: RECEPTOR FOR THE PUTATIVE CYTOKINE 4-1BBL. POSSIBLY			
CC	ACTIVE DURING T CELL ACTIVATION.			
CC	- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	- TISSUE SPECIFICITY: EXPRESSED ON THE SURFACE OF ACTIVATED T			
CC	CELLS.			
CC	- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.			
DR	EMBL; U03397; G571321; -.			
DR	EMBL; L12964; G292238; -.			
DR	PROSITE; PS00652; TNFR_NGFR.			
KW	RECEPTOR; GLYCOPROTEIN; REPEAT; SIGNAL; TRANSMEMBRANE.			
FT	SIGNAL	1	17	POTENTIAL.
FT	CHAIN	18	255	4-1BB LIGAND RECEPTOR.
FT	DOMAIN	18	186	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	187	213	POTENTIAL.
FT	DOMAIN	214	255	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	47	159	3 X TNFR-CYS.
FT	REPEAT	47	86	TNFR-CYS 1.
FT	REPEAT	87	118	TNFR-CYS 2.
FT	REPEAT	119	159	TNFR-CYS 3.
FT	CARBOHYD	138	138	POTENTIAL.
FT	CARBOHYD	149	149	POTENTIAL.

```

FT CONFLICT 107 107 K -> R (IN REF. 2).
SQ SEQUENCE 255 AA; 27899 MW; 331C1176 CRC32;
Query Match 100.0%; Score 1925; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mgnscynivaatl1llvnefrtsl0dpcsnpcagfcdnnrnqicspcpnfsfsgagqr 60
QY 1 MGNSCYNIVAATLLVNEFRTRSL0DPCSNCPAGTFCDDNNRNQICSPCPNFSFSGAGQR 60
Db 61 tcdicrckgkyfrtrkessstnaecdtprghclgagcsmcedckogdeltrkckdc 120
QY 61 TCDICRCKGKYFRTRKESSTNAECDCPRGHCLGAGCSMCEDCOKOGDELTRKCKDC 120
Db 121 cftgndqgkyrcprwtncsl0dgrsvlkgtkedvvcgppvvsfspstlsvtpeagp 180
QY 121 CFTGFNDQKRGICRPWTNCSLDGKSVLVNGTKERDVVCGPSPADLSFGASVTPAPAPAR 180
Db 181 pghspqilseffaltstallflfllrfsyvkrgirklllyfkqfmrpyqtgeedg 240
QY 181 PGHSPQILSEFFALTSTALLFLFLFRFSYVKGRRKLLYFKQFMRPVQTTQEEBDG 240
Db 241 cscrfpeeeeggccl 255
QY 241 CSCRFPEEEEGGCEL 255

RESULT 2
ID 41BB_MOUSE STANDARD; PRT; 256 AA.
AC P20334;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DE 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE 4-1BB LIGAND RECEPTOR PRECURSOR (T-CELL ANTIGEN 4-1BB).
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BAB/C;
RX MEDLINE; 89184547.
RA KWON B.S.; KOZAK C.A.; KIM K.K.; PICKARD R.T.;
RA PROC. NATL. ACAD. SCI. U.S.A. 86:1963-1967(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BAB/C;
RX MEDLINE; 94179805.
RA KWON B.S.; KOZAK C.A.; KIM K.K.; PICKARD R.T.;
RA J. IMMUNOL. 152:2256-2262(1994).
RN [3]
RP CHARACTERIZATION, AND SEQUENCE OF 25-29.
RA MEDLINE; 9319510.
RA POLLOK K.E.; KIM Y.-J.; ZHOU Z.; HURTADO J.; KIN K.K.; PICKARD R.T.;
RA KWON B.S.;
RA J. IMMUNOL. 150:771-781(1993).
RN [4]
RP FUNCTION: RECEPTOR FOR THE PUTATIVE CYTOKINE 4-1BBL. POSSIBLY
ACTIVE DURING T CELL ACTIVATION.
CC -1- SUBUNIT: PRINCIPALLY AN HOMODIMER, BUT ALSO FOUND AS A MONOMER.
CC -1- SUBUNIT: ASSOCIATES WITH P56-LCK.
CC -1- INDUCTION: OPTIMAL BY PMA AND IONOMYCIN.
CC -1- TISSUE SPECIFICITY: EXPRESSED ON THE SURFACE OF ACTIVATED T-CELLS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; J04492; G401122; -.
DR EMBL; U02567; G409178; -.
DR PIR; B32393; B32393.
DR PROSITE; PS00652; TNFR_NGFR.
KW RECEPTOR; GLYCOPROTEIN; REPEAT; SIGNAL; TRANSMEMBRANE.
FT SIGNAL 1 24
FT CHAIN 25 256
FT DOMAIN 25 187 4-1BB LIGAND RECEPTOR.
FT TRANSEM 188 208 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 209 256 CYTOPLASMIC (POTENTIAL).

```

```

FT DOMAIN 17 159 4 X TNFR-CYS.
FT REPEAT 17 45
FT REPEAT 46 85 TNFR-CYS 1.
FT REPEAT 86 117 TNFR-CYS 2.
FT REPEAT 118 159 TNFR-CYS 3.
FT CARBOHYD 128 128 TNFR-CYS 4.
FT CARBOHYD 138 138 POTENTIAL.
SQ SEQUENCE 256 AA; 27598 MW; 5C2352FC CRC32;
Query Match 58.0%; Score 1116; DB 1; Length 256;
Best Local Similarity 57.3%; Pred. No. 6.16e-200;
Matches 145; Conservative 49; Mismatches 55; Indels 4; Gaps 4;

Db 1 mgnscynivaatl1llvnefrtsl0dpcsnpcagfcdnnrnqicspcpnfsfsgagqr 60
QY 1 MGNSCYNIVAATLLVNEFRTRSL0DPCSNCPAGTFCDDNNRNQICSPCPNFSFSGAGQR 60
Db 60 ncnlcrvaagfyfrfkfesshnaecceiegfhclgppctcckdcprgageltkxcktc 119
QY 61 TCDICRCKGKYFRTRKESSTNAECDCPRGHCLGAGCSMCEDCOKOGDELTRKCKDC 120
Db 120 slgtfndqgkyrcprwtncsl0dgrsvlkgtkedvvcgppvvsfspstlsvtpeagp 179
QY 121 CFTGFNDQKRGICRPWTNCSLDGKSVLVNGTKERDVVCGPSPADLSFGASVTPAPAPAR 179
Db 180 g-ghslqyltflalts-allaiflflfsvlkwirkkphlfpkfttaageed 237
QY 180 PGHSPQILSEFFALTSTALLFLFLFRFSYVKGRRKLLYFKQFMRPVQTTQEEBDG 239
Db 238 acscrfpeeeeggccl 250
QY 240 GCSCRFPEEEEGGCEL 252

RESULT 3
ID CD27_MOUSE STANDARD; PRT; 250 AA.
AC P41272;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE CD27L RECEPTOR PRECURSOR (T-CELL ACTIVATION ANTIGEN CD27).
GN CD27.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B6/CBAF1; TISSUE-THYMUS;
RX MEDLINE; 93209286.
RA GRAVESTON L.A.; BLOM B.; NOITEN L.A.; DE VRIES E.; VAN DER HORST G.;
RA OSSENDORP F.; BORST J.; LOENEN W.A.;
RA EUR. J. IMMUNOL. 23:943-950(1993).
RN [2]
RP FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD27L. MAY PLAY
A ROLE IN SURVIVAL OF ACTIVATED T-CELLS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: IN THYMUS AND SPLEEN, BUT NOT IN NON-LYMPHOID
TISSUES.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; L24495; -. NOT_ANNOTATED_CDS.
DR PIR; A49053; A49053.
DR PROSITE; PS00652; TNFR_NGFR.
KW T-CELL; GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL; RECEPTOR; REPEAT.
FT SIGNAL 1 20
FT CHAIN 21 250
FT DOMAIN 21 182 CD27L RECEPTOR.
FT TRANSEM 183 203 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 204 250 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 204 250 3 X TNFR-CYS.
FT REPEAT 26 63 TNFR-CYS 1.
FT REPEAT 64 104 TNFR-CYS 2.
FT REPEAT 105 141 TNFR-CYS 3.
FT CARBOHYD 95 95 POTENTIAL.

```

FT	CARBONYD	162	162	POTENTIAL.
SO	SEQUENCE	250 AA;	28164 MM;	AF6F74ED CAC32;
Query Match				
Best Local Similarity		10.7%;	Score 206;	DB 2; Length 250;
Matches		26; Conservative	14; Mismatches	26; Indels 4; Gaps 4;
Db	53	ceqdataagcdpcipgtsfspdnyhtphcecrhcnsgfilrn-clvtanaescskmg	111	
QY	37	CDNNRNOI-CSPC-PNNSFSSAGGQRT-CDICROCKGVFTFRKCSSTNAECSDCTPGF	93	
Db	112	crdcectcd	121	
QY	94	CLGAGCSKCE	103	
RESULT	4	STANDARD;	PRT;	260 AA.
ID	CD27_HUMAN			
AC	P26842;			
DT	01-AUG-1992 (REL. 23, CREATED)			
DI	01-FEB-1992 (REL. 23, LAST SEQUENCE UPDATE)			
DE	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)			
GN	CD27L RECEPTOR PRECURSOR (T-CELL ACTIVATION ANTIGEN CD27) (T14).			
GC	CD27.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUDAROTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
CC	EUTHERIA; PRIMATES.			
LN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-MONONUCLEAR BLOOD CELLS;			
RA	MEDLINE; 92013149.			
FX	CAMERINI D., WALZ G., LOENEN W.A.M., BORST J., SEED B.,			
RL	J. IMMUNOL. 147:3165-3169(1991).			
LN	[2]			
RP	SEQUENCE FROM N.A.			
RC	MEDLINE; 93094588.			
RA	LOENEN W.A., GRAVESTEIN L.A., BEUMER S., MELIEF C.J., HAGEMELDER A.,			
RL	BORST J.;			
CC	J. IMMUNOL. 149:3937-3943(1992).			
CC	-1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD27L. MAY PLAY			
CC	A ROLE IN SURVIVAL OF ACTIVATED T-CELLS.			
CC	-1- SUBUNIT: HOMODIMER.			
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-1- TISSUE SPECIFICITY: FOUND IN MOST T LYMPHOCYTES.			
CC	-1- PM: PHOSPHORYLATED AND O-GLYCOSYLATED (PROBABLE).			
CC	-1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.			
DR	EMBL; M63928; A46085; -			
PIR	A46454; A46454.			
PIR	A46517; A46517.			
DR	MIM; 186711; -			
DR	PROSITE; PS00652; TNFR_NGFR.			
KW	T-CELL; GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL; RECEPTOR; REPEAT;			
KW	PHOSPHORYLATION.			
FT	SIGNAL	1	20	POTENTIAL.
FT	CHAIN	21	260	CD27L RECEPTOR.
FT	DOMAIN	21	191	EXTRACELLULAR (POTENTIAL).
FT	TRANSEM	192	212	POTENTIAL.
FT	DOMAIN	213	260	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	26	141	3 X TNFR-CYS.
FT	REPEAT	26	63	TNFR-CYS 1.
FT	REPEAT	64	104	TNFR-CYS 2.
FT	REPEAT	105	141	TNFR-CYS 3.
FT	CARBONYD	95	95	POTENTIAL.
FT	MOD RES	219	219	PHOSPHORYLATED (POTENTIAL).
FT	CONFLICT	59	59	A -> T (IN REF. 2).
SO	SEQUENCE	260 AA;	29156 MM;	F707E40F CRC32;
Query Match				
Best Local Similarity		10.3%;	Score 199;	DB 2; Length 260;
Matches		27; Conservative	14; Mismatches	25; Indels 4; Gaps 4;
Db	53	cdghnkaagcdpcipgtsfspdnyhtphcecrhcnsgfilrn-clvtanaescskmg	111	
QY	37	CDNNRNOI-CSPC-PNNSFSSAGGQRT-CDICROCKGVFTFRKCSSTNAECSDCTPGF	93	

[illegible]

RA FORCE W.R., WILLIAMS-ABBOTT L., BROWNING J., HESSION C., TIZARD R.,
 RA WARE C.F.;
 RA SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE; 96163885.
 RA NAKAMURA T., TASHIRO K., NAZAREA M., NAKANO T., SASAYAMA S.,
 RA HONJO T.;
 RA GENOMICS 30:312-319(1995).
 RL -1- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
 CC IMMUNE DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC EMBL; U29173; G881621; -.
 DR EMBL; L38423; G600223; -.
 DR EMBL; U30798; G1061327; -.
 KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL.
 FT SIGNAL 1 30
 FT CHAIN 31 415
 FT DOMAIN 31 223
 FT TRANSMEM 224 244
 FT DOMAIN 245 415
 FT DOMAIN 42 213
 FT REPEAT 42 81
 FT REPEAT 82 124
 FT REPEAT 125 170
 FT REPEAT 171 213
 FT DISULFID 43 58
 FT DISULFID 59 72
 FT DISULFID 62 80
 FT DISULFID 83 98
 FT DISULFID 101 116
 FT DISULFID 104 124
 FT DISULFID 126 132
 FT DISULFID 139 150
 FT DISULFID 142 169
 FT DISULFID 172 187
 FT CARBOHYD 40 40
 FT CARBOHYD 179 179
 FT SEQUENCE 415 AA; 44956 MW; 3C5DD121 CRC32;
 Query Match 9.5%; Score 183; DB 9; Length 415;
 Best Local Similarity 35.5%; Pred. No. 8,48e-14;
 Matches 33; Conservative 17; Mismatches 33; Indels 10; Gaps 6;
 Db 53 pmhvcscrcpgpgefvvcscrsqdvctvcktophnsyehnmhstcqlcrpdcvlgfee 112
 QY 23 SLOPBC-SNCPAG--TF--CDNNRNQICSPCPNFSFSSAGGQ-RTCDICRCKGV--FRT 74
 113 vapctsdtkacrcqpgmcyvldncvchceee 145
 QY 75 RKECSSTSNACDDCTPGFHC--LGAGCSMCEPD 105
 RESULT 7
 ID VC22-VARV STANDARD; PRT: 349 AA.
 AC P34015;
 DT 01-FEB-1994 (REL. 28, CREATED)
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
 DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
 DE PROTEIN C22/B28 HOMOLOG.
 GN GAR.
 OS VARIOLOA VIRUS.
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXYVIRIDAE; CHORDOPOXYVIRINAE;
 CC ORTHOPOXYVIRUSES.
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-INDIA-1967 / ISOLATE IND3;
 RX MEDLINE; 93202281.
 RA SHEHELEKUNOV S.N., BLINOV V.M., SANDAKHCHIEV L.S.;
 RL FEBS LETT. 319:80-83(1993).
 CC -1- SIMILARITY: CONTAINS TWO LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGIONS.
 DR EMBL; X69198; G457087; -.
 Query Match 9.0%; Score 174; DB 10; Length 325;
 Best Local Similarity 31.3%; Pred. No. 2,63e-12;
 Matches 25; Conservative 15; Mismatches 35; Indels 5; Gaps 4;

DR EMBL; X67117; G516449; -.
 DR PIR; D36858; D36858.
 DR PIR; S35987; S35987.
 DR PIR; S46888; S46888.
 DR HSSP; P19438; TNFR.
 DR PROSITE; PS00652; TNFR_NGFR.
 KW REPEAT.
 FT DOMAIN 31 108
 FT REPEAT 31 66
 FT REPEAT 67 108
 FT SEQUENCE 349 AA; 38189 MW; 50DDB435 CRC32;
 Query Match 9.3%; Score 179; DB 10; Length 349;
 Best Local Similarity 34.1%; Pred. No. 3,93e-13;
 Matches 29; Conservative 20; Mismatches 29; Indels 7; Gaps 4;
 Db 44 clscppgyaarlcdskntctpcsgyftsrnhlpaciscngrcnsqvetrsctt 103
 QY 28 CSNCPAGTF---CDNNRNQICSPCPNFSFSSAGGQ-RTCDICR-CKGVFRTKCSST 81
 Db 104 hnricecpgyycllkgsagc-kac 127
 QY 82 SNABCDCTPGFHC LGAGCSMCEDC 106
 RESULT 8
 ID VT2-SEVKA STANDARD; PRT: 325 AA.
 AC P25943;
 DT 01-MAY-1992 (REL. 22, CREATED)
 DT 01-FEB-1992 (REL. 22, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN T2).
 GN T2.
 OS SHOPE FIBROMA VIRUS (STRAIN KASZA) (SFV).
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXYVIRIDAE; CHORDOPOXYVIRINAE;
 CC LEPORIPOVIRUSES.
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE; 87321103.
 RA UPTON C., DELANGE A.M., MCFADDEN G.;
 RL VIROLOGY 160:20-30(1987).
 RN [2]
 RN FUNCTION.
 RX MEDLINE; 91207415.
 RA SMITH C.A., DAVIS T., WIGNALL J.M., DIN W.S., FARAH T., UPTON C.,
 RA MCFADDEN G., GOODWIN R.G.;
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 176:335-342(1991).
 CC -1- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO
 CC REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL
 CC ANTIYIRAL EFFECTS OF THE CYTOKINE.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC EMBL; M17433; -. NOT_ANNOTATED_CDS.
 DR EMBL; A23727; E199408; -.
 DR PIR; B43692; B43692.
 DR HSSP; P19438; TNFR.
 DR PROSITE; PS00652; TNFR_NGFR.
 KW RECEPTOR; GLYCOPROTEIN; REPEAT; SIGNAL.
 FT SIGNAL 1 16
 FT CHAIN 17 325
 FT DOMAIN 17 186
 FT REPEAT 27 62
 FT REPEAT 63 104
 FT REPEAT 105 147
 FT REPEAT 148 186
 FT CARBOHYD 105 105
 FT CARBOHYD 181 181
 FT CARBOHYD 205 205
 FT CARBOHYD 238 238
 FT SEQUENCE 325 AA; 35132 MW; C9D2C87B CRC32;
 Query Match 9.0%; Score 174; DB 10; Length 325;
 Best Local Similarity 31.3%; Pred. No. 2,63e-12;
 Matches 25; Conservative 15; Mismatches 35; Indels 5; Gaps 4;

```
Db 49 asrlcpgsntvcspecdyftfctasthnapacvscrcpctghlscspcdtrthrvencst 108
QY 33 AGFCNNRNQICSPCPNPFSSAGGQ-RFTCDICR-QCKGVFTRRRCSSSTNAECDCPT 90
Db 109 gnycllkqgngcrccapctk 128
QY 91 GFHCL--G-AGCSMCBODCK 107

RESULT 9
ID CD40_MOUSE STANDARD: PRT: 289 AA.
AC P27512:
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDW40).
GN CD40.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92105763.
RA TORES R.M., CLARK E.A.;
RL J. IMMUNOL. 148:620-626(1992).
RN [2]
RP REVISIONS.
RC STRAIN-BALB/C;
RA TORES R.M.;
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-LIVER;
RX MEDLINE: 93094586.
RA GRIMALDI J.C., TORES R., KOZAK C.A., CHANG R., CLARK E.A.,
RL J. IMMUNOL. 149:3921-3926(1992).
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: M83312; G1553059; -.
DR EMBL: M94126; G192526; -.
DR EMBL: M94129; G192526; JOINED.
DR EMBL: M94128; G192526; JOINED.
DR EMBL: M94127; G192526; JOINED.
DR PIR: A46476; A46476.
DR HSSP: P19438; 1TNFR.
DR PROSITE: PS00652; TNFR_NGFR.
KM RECEPTOR; B-CELL; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT; SIGNAL.
FT SIGNAL. 1 19
FT CHAIN. 20 289
FT DOMAIN. 20 193 - POTENTIAL.
FT TRANSMEM 194 215 EXTRACELLULAR (POTENTIAL).
FT DOMAIN. 216 289 CYTOPLASMIC (POTENTIAL).
FT DOMAIN. 25 187 4 X TNFR-CYS.
FT REPEAT. 25 60 TNFR-CYS 1.
FT REPEAT. 61 103 TNFR-CYS 2.
FT REPEAT. 104 144 TNFR-CYS 3.
FT REPEAT. 145 187 TNFR-CYS 4.
FT CARBOHYD 153 153 POTENTIAL.
SQ SEQUENCE 289 AA; 32111 MW; D8D70A2C CRC32;

Query Match 8.9%; Score 171; DB 2; Length 289;
Best Local Similarity 31.2%; Pred. No. 8,18e-12;
Matches 43; Conservative 21; Mismatches 58; Indels 16; Gaps 11;
```

```
QY 93 HCLGACSCMEOD--CKG---QEL-T-KKG--CKDCGTFNDQR--GICRPTNCSL 141
Db 170 knlevlqkgtstgnvieg 187
QY 142 DGKSVLVNGTKERDVYCG 159

RESULT 10
ID OX40_RAT STANDARD: PRT: 271 AA.
AC P15725;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN) (MRC OX40).
GN TXGPII OR OX40.
OS EUTHERIA NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-T-CELL;
RX MEDLINE: 90214614.
RA MALETT S., FOSSUM S., BARCLAY A.N.;
RL EMBO J. 9:1063-1068(1990).
CC -1- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: ACTIVATED T-CELLS.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: X17037; G57831; -.
DR PIR: S08036; S08036.
DR PIR: S12783; S12783.
DR PROSITE: PS00652; TNFR_NGFR.
KM RECEPTOR; T-CELL; ANTIGEN; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT;
RN SIGNAL.
FT SIGNAL. 1 19
FT CHAIN. 20 271
FT DOMAIN. 20 210 OX40L RECEPTOR.
FT TRANSMEM 211 235 EXTRACELLULAR (POTENTIAL).
FT DOMAIN. 236 271 POTENTIAL.
FT DOMAIN. 25 164 CYTOPLASMIC (POTENTIAL).
FT DOMAIN. 25 60 4 X TNFR-CYS.
FT REPEAT. 25 60 TNFR-CYS 1.
FT REPEAT. 61 102 TNFR-CYS 2.
FT REPEAT. 103 123 TNFR-CYS 3 (INCOMPLETE).
FT REPEAT. 124 164 TNFR-CYS 4.
FT CARBOHYD 143 143 POTENTIAL.
SQ SEQUENCE 271 AA; 29895 MW; 65460ED2 CRC32;

Query Match 8.6%; Score 166; DB 7; Length 271;
Best Local Similarity 30.6%; Pred. No. 5,33e-11;
Matches 38; Conservative 25; Mismatches 48; Indels 13; Gaps 5;
```

```
Db 51 ctalektg-chpodsgefaqwnreirchqhncpuglrvkkgtaesdtyctckegq 109
QY 37 CDN-NNNQICSPCPNPFSSA-GGORTCDICRQCK-G-VRTRRRCSSSTNAECDCPT 92
Db 110 hctskceacqhtpctpgfygmatalettctvchpccpygffinsgslfckypwtscd 169
```

```
RESULT 11
ID TNRC_HUMAN STANDARD: PRT: 435 AA.
AC P36941;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR (TUMOR NECROSIS FACTOR RECEPTOR
DE 2 RELATED PROTEIN) (TUMOR NECROSIS FACTOR C RECEPTOR).
GN TNFR.
```

```

OS  HOMO SAPIENS (HUMAN).
OC  EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC  EUTHERIA; PRIMATES.
RN  (1)
RP  SEQUENCE FROM N.A.
RC  TISSUE=LIVER;
RX  MEDLINE: 93252381.
RA  BAENS M., CHAFANET M., CASSIMAN J.J., DEN BERGHE H., MARYENEN P.,
RL  GENOMICS 16:214-218(1993).
RN  (2)
RP  FUNCTION.
RX  MEDLINE: 94225209.
RA  CROME P.D., VAN ARSDALE T.L., WALTER B.N., WARE C.F., HESSION C.,
RC  EHRENFELS B., BROWNING J.L., DIN W.S., GOODWIN R.G., SMITH C.A.;
RL  SCIENCE 264:707-710(1994).
CC  -1- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
CC  IMMUNE DEVELOPMENT.
CC  -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC  -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR  EMBL: L04270; G339762; -.
DR  FIM: 600979; -.
DR  HSSP: P19999; 1CLG.
DM  PROSITE: PS00652; TNFR_NGFR.
KW  RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL.
FT  SIGNAL 1 30 POTENTIAL.
FT  CHAIN 31 435 LYMPHOTOXIN-BETA RECEPTOR.
FT  DOMAIN 31 227 EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM 228 248 POTENTIAL.
FT  DOMAIN 249 435 CYTOPLASMIC (POTENTIAL).
FT  DOMAIN 42 211 4 X TNFR-CYS.
FT  REPEAT 42 81 TNFR-CYS 1.
FT  REPEAT 82 124 TNFR-CYS 2.
FT  REPEAT 125 168 TNFR-CYS 3.
FT  REPEAT 169 211 TNFR-CYS 4.
FT  DISULFID 43 58 BY SIMILARITY.
FT  DISULFID 59 72 BY SIMILARITY.
FT  DISULFID 62 80 BY SIMILARITY.
FT  DISULFID 83 98 BY SIMILARITY.
FT  DISULFID 101 116 BY SIMILARITY.
FT  DISULFID 104 124 BY SIMILARITY.
FT  DISULFID 126 132 BY SIMILARITY.
FT  DISULFID 139 148 BY SIMILARITY.
FT  DISULFID 142 167 BY SIMILARITY.
FT  DISULFID 170 185 BY SIMILARITY.
FT  CARBOHYD 40 40 POTENTIAL.
FT  CARBOHYD 177 177 POTENTIAL.
SQ  SEQUENCE 435 AA; 46709 MW; 203882DD CRC32;

Query Match 8.2%; Score 158; DB 9; Length 435;
Best Local Similarity 31.5%; Pred. No. 1.03e-09;
Matches 52; Conservative 26; Mismatches 66; Indels 21; Gaps 13;

Db 59 csrccpptyvskscrlrdtvcatacensynehmyltlclqclrcpbymllelaptcs 118
OY 28 CSNCPAGTF---CDNNRNQICSPCPNFSFA-GGORTDRCROCKGV--FTRKRCSS 80
DB 119 krttgcrcqpmfcaalectccllscdcppteaeikdevykgmhcyckaghifnt 178
OY 81 TSNAECDCTPPGFHCLG-A-GCSMCE--QDCKQG-Q-ELTK--KGCKDC--C-FGTF-ND 127
DB 179 ssparscphtcengglveaapgtasdtlcknplelplpemsq 223
OY 128 QKRGT-CRPWTNCSLDGKSVLVNGTKERDVYCGPSPADLSFGASS 171

```

```

OS  MUS MUSCULUS (MOUSE).
OC  EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC  EUTHERIA; RODENTIA.
RN  (1)
RP  SEQUENCE FROM N.A.
RC  STRAIN=BALB/C;
RX  MEDLINE: 94044750.
RA  CALDERHEAD D.M., BUHLMANN J.E., VAN DEN EERTWEGH A.J.,
RL  J. IMMUNOL. 151:5261-5271(1993).
RN  (2)
RP  SEQUENCE FROM N.A.
RX  MEDLINE: 95255413.
RA  BIRKELAND M.L., COPELAND N.G., GILBERT D.J., JENKINS N.A.,
RL  BARCLAY A.N.;
RL  EUR. J. IMMUNOL. 25:926-930(1995).
CC  -1- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC  -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC  -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR  EMBL: Z21674; G312828; -.
DR  EMBL: X85214; G732819; -.
KW  RECEPTOR; T-CELL; ANTIGEN; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT;
KW  SIGNAL.
FT  SIGNAL 1 19 POTENTIAL.
FT  CHAIN 20 272 OX40L RECEPTOR.
FT  DOMAIN 20 211 EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM 212 236 POTENTIAL.
FT  DOMAIN 237 272 CYTOPLASMIC (POTENTIAL).
FT  DOMAIN 26 165 4 X TNFR-CYS.
FT  REPEAT 26 61 TNFR-CYS 1.
FT  REPEAT 62 103 TNFR-CYS 2.
FT  REPEAT 104 124 TNFR-CYS 3 (INCOMPLETE).
FT  REPEAT 125 165 TNFR-CYS 4.
FT  CARBOHYD 144 144 POTENTIAL.
FT  CARBOHYD 15 15 A -> G (IN REF. 2).
SQ  SEQUENCE 272 AA; 30153 MW; 67D1B978 CRC32;

Query Match 7.3%; Score 140; DB 7; Length 272;
Best Local Similarity 30.6%; Pred. No. 6.53e-07;
Matches 38; Conservative 21; Mismatches 52; Indels 13; Gaps 5;

Db 52 cdhttrdtlhpcegtfyneavnydtcgtcgmhrgselskngncptqdtvcrprgtg- 110
OY 37 CDNNRNQICSPCPNFSFSSAGGORTDRCG-KVFTRKRCSSTSNAEDCTPPGFHC 94
DB 111 -----pr-qds--gyklygv-dcvpcpghfispnngackpwtnctlsqktrhpsdsi 160
OY 95 LGAGCSMCQDCKQGGELTKKGCDCFCFTFDQKRGICRPWTNCSLDGKSVLVNGTKER 154
DB 161 dvvc 164
OY 155 DVVC 158

```

```

RESULT 12
ID OX40_MOUSE STANDARD: PRT: 272 AA.
AC P47741:
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN).
GN TXGP1L OR OX40.

```

```

RESULT 13
ID OX40_HUMAN STANDARD: PRT: 277 AA.
AC P43489:
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE OX40L RECEPTOR PRECURSOR (ACT35 ANTIGEN) (TXA-TRANSCRIPTIONALLY
DE ACTIVATED GLYCOPROTEIN 1 RECEPTOR).
GN TXGP1L.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE: 94170844.
RA LATZA U., DUERKOP H., SCHNITZER S., RINGELING J., EITELBACH F.,
RA HUMMEL M., FORATSCHE C., STEIN H.;
RL EUR. J. IMMUNOL. 24:677-683(1994).
CC -1- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.

```


CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL: X75962; G472958; -.
 DR MIM: 600315; -.
 DR PROSITE: PS00652; TNFR_NGFR.
 KW RECEPTOR; T-CELL; ANTIGEN; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT;
 KM SIGNAL.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 277 OX40L RECEPTOR.
 FT DOMAIN 29 214 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 215 235 POTENTIAL.
 FT DOMAIN 236 277 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 30 169 4 X TNFR-CYS.
 FT REPEAT 30 65 TNFR-CYS 1.
 FT REPEAT 66 107 TNFR-CYS 2.
 FT REPEAT 108 126 TNFR-CYS 3 (INCOMPLETE).
 FT REPEAT 127 167 TNFR-CYS 4.
 FT CARBOHYD 146 146 POTENTIAL.
 FT CARBOHYD 160 160 POTENTIAL.
 SQ SEQUENCE 277 AA; E0E6765 CRC32:
 Query Match 7.3%; Score 141; DB 7; Length 277;
 Best Local Similarity 26.4%; Pred. No. 4,60e-07;
 Matches 38; Conservative 29; Mismatches 61; Indels 16; Gaps 9;
 Db 56 csgtqvtcpgpgfyndvsvskpckpctwcnlrsgerkqlctatqdvrc-ra--- 111
 QY 37 CDNNRNQICSPCPNFSAGGQRTCDICRCK-GVFRTKRE-CSTSNMECDCTPGFHC 94
 Db 112 -gtq-pl-d-syhpqv-d-----capcpghfispdndgackpwhtclaghtlqpaans 162
 QY 95 LGAGCSKCEODCKGGGLTKGCKCCGTFPNDKRGICRWNTCSJDGSAVLVTKER 154
 Db 163 dalcedrdpatqpgelqpparp 186
 QY 155 DVCV-GSPPADLSRGASVTPPAR 177
 RESULT 14
 ID NGFR_RAT STANDARD; PRT; 425 AA.
 AC P07174;
 DT 01-APR-1988 (REL. 07, CREATED)
 DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE LOW-AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (NGF RECEPTOR)
 DE (GP80-LINGFR).
 GN NGFR.
 OS RATRUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 87115859.
 RA RADEKE M.J., MISKO T.P., HSU C., HERZENBERG L.A., SHOOTER E.M.;
 RL NATURE 325:593-597(1987).
 RN [2]
 RP SEQUENCE OF 1-22 FROM N.A.
 RC TISSUE-LIVER;
 RX MEDLINE: 93077038.
 RA METSIS M., TIMUSK T., ALLIKMETS R., SAARMA M., PERSSON H.;
 RL GENE 121:247-254(1992).
 CC -1- FUNCTION: LOW AFFINITY RECEPTOR WHICH CAN BIND TO NGF, BDNF,
 CC NT-3, AND NT-4.
 CC -1- SUBUNIT: NGF RECEPTOR CAN FORM A HOMODIMER THROUGH DISULFIDE
 CC BOND FORMATION.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- PTM: N- AND O-GLYCOSYLATED AND IS PHOSPHORYLATED ON SERINE.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL: X6137; G56756; -.
 DR EMBL: X61269; -; NOT_ANNOTATED_CDS.
 DR PIR: A26431; A26431.
 DR HSSP: P19438; INTR.
 DR PROSITE: PS00652; TNFR_NGFR.

DR PROSITE: PS0017; DEATH DOMAIN.
 KW RECEPTOR; NEUROGENESIS; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT;
 KM PHOSPHORYLATION; SIGNAL.
 FT SIGNAL 1 29 NGF RECEPTOR.
 FT CHAIN 30 425 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 30 251 POTENTIAL.
 FT TRANSMEM 252 273 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 274 425 4 X TNFR-CYS.
 FT DOMAIN 32 190 TNFR-CYS 1.
 FT REPEAT 32 66 TNFR-CYS 2.
 FT REPEAT 67 108 TNFR-CYS 3.
 FT REPEAT 109 148 TNFR-CYS 4.
 FT REPEAT 149 190 SER/THR-RICH.
 FT DOMAIN 198 249 DEATH DOMAIN.
 FT DOMAIN 354 419 BY SIMILARITY.
 FT DISULFID 33 44 BY SIMILARITY.
 FT DISULFID 45 58 BY SIMILARITY.
 FT DISULFID 48 65 BY SIMILARITY.
 FT DISULFID 68 84 BY SIMILARITY.
 FT DISULFID 87 100 BY SIMILARITY.
 FT DISULFID 90 108 BY SIMILARITY.
 FT DISULFID 110 123 BY SIMILARITY.
 FT DISULFID 126 139 BY SIMILARITY.
 FT DISULFID 129 147 BY SIMILARITY.
 FT DISULFID 150 165 BY SIMILARITY.
 FT DISULFID 168 181 BY SIMILARITY.
 FT DISULFID 171 189 BY SIMILARITY.
 FT CARBOHYD 61 61 POTENTIAL.
 FT CARBOHYD 71 71 POTENTIAL.
 SQ SEQUENCE 425 AA; 45432 MM; 7D78F258 CRC32:
 Query Match 7.2%; Score 138; DB 6; Length 425;
 Best Local Similarity 31.0%; Pred. No. 1,31e-06;
 Matches 22; Conservative 18; Mismatches 26; Indels 5; Gaps 3;
 Db 124 eacsvcevgsglvtfcgdkntvceepgtydeanhvpcplctvctedterqlrectp 183
 QY 26 DPCSNCPAGT--F-CDNNRNQICSPCPNFSAGGQR-TCDICRCKGVFRTKRCSS 80
 Db 184 wadaeeceipg 194
 QY 81 TSNMECDCTPG 91
 RESULT 15
 ID CD40_HUMAN STANDARD; PRT; 277 AA.
 AC P25942;
 DT 01-MAY-1992 (REL. 22, CREATED)
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
 DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
 DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDM40).
 GN CD40.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 89356608.
 RA STRAMENKOVIC I., CLARK E.A., SEED B.;
 RL EMBO J. 8:1403-1410(1989).
 CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: B-CELLS AND IN PRIMARY CARCINOMAS.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL: X60592; G29851; -.
 DR PIR: S04460; S04460.
 DR MIM: 109335; -.
 DR PROSITE: PS00652; TNFR_NGFR.
 KW RECEPTOR; B-CELL; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT; SIGNAL.
 KM SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 277 CD40L RECEPTOR.
 FT DOMAIN 20 193 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 194 215 POTENTIAL.

FT	DOMAIN	216	277	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	25	187	4 X TNFR-CYS.
FT	REPEAT	25	60	TNFR-CYS 1.
FT	REPEAT	61	103	TNFR-CYS 2.
FT	REPEAT	104	144	TNFR-CYS 3.
FT	REPEAT	145	187	TNFR-CYS 4.
FT	CARBOHYD	153	153	POTENTIAL.
FT	CARBOHYD	180	180	POTENTIAL.
SO	SEQUENCE	277 AA;	30619 MM;	3B284411 CRC32;

Query Match 7.1%; Score 136; DB 2; Length 277;
 Best Local Similarity 28.3%; Pred. No. 2.63e-06;
 Matches 39; Conservative 24; Mismatches 61; Indels 14; Gaps 10;

Db	51	ctetctecclpcgeselfidwrethchqkycdpnlglyvqgkytsetdlctceegwh	110
QY	37	CDNNRNQICSPCPNPSFSSAGGRT-CDICRQK-GV-FRTKKECSSTSMACDCTPGEH	93
Db	111	ctseacescvlhrcspgfygkqiatgvsdtlcepcpygffsnvsafekchpytscetk	170
QY	94	CLGAGCSMCE-Q-DCKQGQELTK--KCKD--C--C-FGTINDQKRGI--CRPWTCSLD	142
Db	171	dlvvqgagtnktdvvcgp	188
QY	143	GKSYLVNNGTKERDVVCGP	160

Search completed: Tue Mar 24 17:04:31 1998
 Job time : 48 secs.